Earth-Science Reviews 193 (2019) 220–236

Contents lists available at ScienceDirect

Earth-Science Reviews

journal homepage: www.elsevier.com/locate/earscirev

Accounting for uncertainty in remotely-sensed measurements of river planform change

Mitchell Donovan[a,*], Patrick Belmont[a], Bastiaan Notebaert[b], Tyrel Coombs[a], Phillip Larson[c], Michael Souffront[a]

[a] Utah State University, Department of Watershed Sciences, Quinney College of Natural Resources, Utah State University, Logan, UT 84321, United States
[b] University of Leuven, Belgium
[c] Department of Geography, Earth Science Programs, AGES Laboratory, Minnesota State University, Mankato, MN 56001, United States

A B S T R A C T

Increased availability and resolution of remotely-sensed (RS) imagery of Earth's surface has greatly enhanced the precision, spatial extent, and temporal frequency at which we can analyze river channel migration and width changes. Despite a body of research identifying and quantifying sources of uncertainty inherent in such data, no framework has emerged to comprehensively quantify and handle uncertainty. Herein, we summarize and evaluate present best practices, test new approaches to quantify and handle uncertainty, and provide recommendations for future work using remotely-sensed measurements of river migration and width changes. While our research focuses on river systems, the principles and approaches are applicable to research delineating boundaries or using boundaries to measure changes: glacier retreat or advance, erosion or deposition along coastlines and lakeshores, changes in wetland extent, expansion or contraction of vegetation (e.g., deforestation), cliff retreat, sea level rise due to climate change, change in aeolian depositional systems, and anthropogenic/political boundary disputes. From our results, the following conclusions and recommendations arise:

1. Planform change measurements should span spatial intervals larger than coherent units of adjustment to avoid spatial autocorrelation.
2. Uncertainty in manual riverbank delineations is dominated by arbitrary user inconsistency rather than poor image quality (i.e., resolution, colour versus grayscale, year of acquisition) or environmental conditions (i.e., shadows and vegetation cover).
3. We recommend that digitizations follow the vegetated boundary that best approximates bankfull width, whenever possible, to avoid inconsistency along ambiguous reaches.
4. Using a spatially variable level of error detection (LoD) threshold improves the quantity and quality of retained measurements relative to a uniform LoD.
5. After applying a LoD threshold, we recommend first using expert discretion to manually classify any 'nondetect' measurements that qualify as 'significant' measurements of zero (i.e., no change actually occurred).
6. Subsequently, three methods may be used for handling the remaining nondetects; Kaplan-Meier (KM) and Maximum Likelihood Estimators (MLE). The specific approach chosen for handling nondetects is contingent upon each case, but can be guided and informed by descriptions and assumptions of each method, references to external resources, and results of our river-focused analyses.
7. Finally, we encourage a focus on improving the simplicity, generalizability, and open-source opportunities of tools and packages used for calculating river planform change and spatially variable uncertainty, thereby enabling a common platform to measure and compare results.

1. Introduction

River channels are among the most dynamic landforms on Earth's surface, sweeping laterally across valley bottoms- often in subtle and sometimes catastrophic- ways over event-, decadal-, and millennial timescales. Remotely-sensed imagery is increasingly used to delineate channel boundaries to measure changes in river planform such as lateral migration (Hickin and Nanson, 1984; Gurnell et al., 1994; Gaeuman et al., 2005b; Lauer and Parker, 2008; Constantine et al., 2014; Donovan et al., 2015, 2016; Morais et al., 2016), as well as channel width (Winterbottom, 2000; Pavelsky and Smith, 2008; Swanson et al., 2011; Downs et al., 2013; Lauer et al., 2017). These measurements provide a basis for understanding effective management strategies for erosion along the riparian corridor (Micheli et al., 2004; Piégay et al., 2005), providing input to sediment budgets (Allmendinger et al., 2007; Belmont et al., 2011; Smith et al., 2011), and automated characterization of single- versus multi-threaded river planforms (Rowland et al., 2016). Such measurements also inform our understanding of important issues including erosional hazards caused by migrating streams (Lawler, 1993; Piégay et al., 1997, 2005; Hughes et al., 2006; Rhoades et al., 2009), understanding the impact of anthropogenic modifications to the fluvial system (Shields et al., 2000;

* Corresponding author.
E-mail address: mdonovan@aggiemail.usu.edu (M. Donovan).

https://doi.org/10.1016/j.earscirev.2019.04.009
Received 17 December 2018; Received in revised form 23 March 2019; Accepted 7 April 2019
Available online 19 April 2019
0012-8252/

Table 1
Factors affecting uncertainty in remotely-sensed images and measurements of planform change.

Factors affecting uncertainty in remotely sensed images and measurements of planform changes			
Type	Expected influencial factors	Considerations	Reference
Georeferencing uncertainty	Image warping (camera optics)	Periphery of images affected relatively more than centers	Fryer and Brown (1986)
	Georeferenced control points	Hard points reduce error compared to soft points	Mount et al. (2003); Hughes et al. (2006)
	Image resolution/quality	Coarser resolution decreases ability to find GCPs	–
	Anthropogenic changes	Urban development changes availability of GCPs	–
	Georectification transformation	2nd-order polynomial is consistently the best choice	Hughes et al. (2006); Donovan et al. (2015)
Digitization (i.e.,delineation) uncertainty	Vegetation density	Vegetation cover reduces visibility of channel edge	Güneralp et al. (2013, 2014); Winterbottom (2000)
	Shadows along river boundary	Shadows reduce visibility of channel edge	Güneralp et al. (2013, 2014)
	Scale of image during delineation	Increased scale (zooming in) reduces delineation error	Liro (2015)
	Water level	Will impact which vegetated-edge boundaries are visible	Lauer et al. (2017), Werbylo et al. (2017)
	Arbitrary user inconsistency	Users are prone to inconsistency in boundary interpretations	Gurnell et al. (1994); This study
	Topographic data availability	High-resolution DEMs increase delineation accuracy	Mount et al. (2013); Donovan et al. (2015)
Automated delineation/ classification uncertainty	Manual vs. [semi]automated	Automated approaches are generally less accurate and vary by which algorithm or software tool is used	Pavelsky and Smith (2008); Rowland et al. (2016); Schwenk et al. (2017)
	Channel width & complexity	Narrow channels and highly complex bar arrangements can reduce automated [and manual] classification accuracy	Rowland et al. (2016); [Werbylo et al., 2017]
Detecting changes in delineated boundaries	Magnitude of changes	Larger changes more likely to exceed level of detection	This study
	Duration between images	Broader timescales span periods of greater change	Donovan and Belmont (2019)
	Reference datum changes	After adjusting images to a common datum, differences may arise due to distinct datum offsets	Anders and Byrnes (1991)
	LoD threshold calculation	Uniform LoD thresholds erroneously reduce number (and quality)* of significant measurements	Lea and Legleiter (2016); *This study

Donovan and Belmont, 2019), and managing riparian habitat (Ward et al., 2002). Ensuring effective management of the river corridor requires that we appropriately quantify and report uncertainty in river migration measurements, lest we run the risk of inappropriately prescribing costly channel and riparian management strategies, including bank stabilization and invasive restoration or rehabilitation practices.

Increased availability and resolution of aerial photography, satellite imagery, unmanned aerial vehicle (UAV) imagery, and LiDAR or digital elevation/terrain models (DEM/DTMs) of Earth's surface have greatly enhanced the precision, spatial extent and temporal frequency with which we can analyze river channel migration (Harpold et al., 2015). Despite an abundance of remotely sensed data and new capabilities enabled by continually evolving software packages, studies of fluvial change based upon remote sensing lack a robust and consistent methodology for quantifying and handling uncertainty (Kiiveri, 1997; Schook et al., 2017; Werbylo et al., 2017). Several studies have provided recent advances to our understanding of uncertainty in measurements of channel width and lateral migration from remotely sensed imagery (Mount et al., 2003; Mount and Louis, 2005; Hughes et al., 2006; Lea and Legleiter, 2016; Werbylo et al., 2017), but no comprehensive framework has been developed. In this way, the methods for measurement of river migration rates lags considerably behind other measurements of topographic change for which rigorous, repeatable and generalizable uncertainty methods have been developed (Brasington et al., 2003; Wheaton et al., 2010; Passalacqua et al., 2015; Schaffrath et al., 2015; Bangen et al., 2016; Vericat et al., 2017; Anderson, 2018).

The goal of this paper is to provide a comprehensive framework for evaluating uncertainty in estimates of river migration and width change by: (1) summarizing relevant research and methods for evaluating uncertainty; (2) highlighting and testing approaches used to estimate channel migration and uncertainty; and (3) filling in gaps regarding how spatial autocorrelations, riparian vegetation, and geomorphic conditions influence uncertainty. This paper does not attempt to address all possible approaches or available tools for analyzing channel migration, but rather discusses the primary considerations and key components of the uncertainty inherent in such measurements. Beyond planform adjustment of river channels, the guidance and results presented herein are applicable to measuring changes in delineated boundaries, including glacier retreat or advance, erosion or deposition along coastlines and lakeshores, changes in wetland extent, expansion or contraction of vegetation (e.g., deforestation), cliff retreat, and political boundary disputes. Furthermore, it will help generate and constrain uncertainty and error estimates for models utilizing such data (Piégay et al., 2005; McMillan and Hu, 2017).

2. Error and uncertainty in geographic information systems

2.1. Background

Measurements of planform change over broad spatial and temporal scales are often derived from series of remotely-sensed images. Such measurements are often made within geographic information systems (GIS) due to their ability to compile and measure spatial (e.g., x, y, z), temporal, and thematic components. Respectively, these three components describe a measurement's location and size within space, the time, speed and/or duration of the measurement, and any associated descriptions or classifying attributes. Measurements of planform change derived from remotely-sensed images contain spatial and temporal components that manifest in magnitudes and rates of change. Such measurements also contain some amount of error, which can be defined as the difference between a measurement of reality, and reality itself (Unwin, 1995; Crosetto and Tarantola, 2001). For planform changes, temporal error is generally absent because the date of image acquisition is usually known. Thematic/classification errors are also irrelevant, leaving only spatial/locational uncertainty in the horizontal and (sometimes) vertical dimensions, to consider. Spatial uncertainty is estimated as the total *possible* error in a given measurement, the

components of which are summarized in Table 1. Estimating the total possible error provides a value with which to quantify a level of detection (LoD) threshold. The LoD is a threshold for determining which measurements are statistically significant. Measurements that fall below the LoD, called 'nondetects', occur due to measurement error exceeding the magnitude of measured change. By quantifying uncertainty alongside measurements of river planform change, we provide transparent and informative data for best management practices along the riparian corridor.

When documenting channel planform and migration change using aerial images, total uncertainty should include uncertainty in image georeferencing and orthorectification, as well as uncertainty in manual- or algorithm-derived channel delineations (Libby et al., 2016). The choice of transformation (e.g., linear, polynomial, kriging, spline, etc.) can have the largest impact on orthorectification uncertainty. A second-order polynomial transformation is recommended for most applications, because it minimizes both image distortion and georeferencing error (Hughes et al., 2006). The quantity, quality, and spatial distribution of georeferenced control points (GCPs) are key factors influencing georeferencing uncertainty, summarized in Table 1 (Lea and Legleiter, 2016). Delineation error has received considerably less attention in the literature. Gurnell et al. (1994) quantified digitization error for a single-threaded highly sinuous channel using average offset of repeated streambank digitizations on 1:10,000 scale maps. They found an average of ± 2 m offset over 18 river km, but were unable to evaluate how error varied spatially or may be affected by overhanging vegetation or shadows, similar to semi- and fully-automated channel delineations (Güneralp et al., 2013, 2014). Working in a mix of braided and anastomosing morphologies, Werbylo et al., 2017 found that while digitizations of multiple users resulted in no significant differences on average channel width, at-a-section differences in width varied by as much as 37 m. The authors conclude that digitizations are more consistent for imagery obtained at high resolution and that flow conditions are the most significant factor impacting error and inconsistency in delineations, with higher flows providing more consistent delineations.

Accurately estimating uncertainty is not only essential for filtering out unreliable measurements, but also for retaining reliable measurements. The latter is particularly relevant for short-term measurements of smaller changes that are rendered obsolete if uncertainty is overestimated (Liro, 2015; Lea and Legleiter, 2016; Donovan and Belmont, 2019). However, when uncertainty cannot be properly quantified, results should be constrained using upper and lower bounds of uncertainty (Kiiveri, 1997; Crosetto and Tarantola, 2001; Donovan et al., 2015; Passalacqua et al., 2015; Lauer et al., 2017) or simply highlighting locations where the measurement is more or less a reflection of noise (i.e., random variability) to provide an estimate of reliability for end users. Using probability and fuzzy positional boundaries has been proposed as a generic approach to estimate probabilistic positional uncertainty in GIS (Kiiveri, 1997; Wheaton et al., 2010). For any application, appropriately evaluating and disclosing uncertainty will improve the quality of results and subsequent applications.

2.2. Techniques and developments in quantifying uncertainty

2.2.1. Georeferencing uncertainty

Georeferencing is the process of placing scanned aerial photographs onto a coordinate plane using known feature locations, referred to as georeferenced control points (GCPs). Optimally, and most often, GCPs are derived from orthorectified images to support maximal accuracy. However, large errors still exist in georeferenced images as a result of errors in the GCPs, associated images, and geometric distortion from cameras/sensors (Fryer and Brown, 1986), scanners, or varying relief. These errors directly affect insights derived from research using image-based delineations for change detection. An estimate of georeferencing uncertainty (i.e., possible X-Y positional error) reflects offset between points on the image and their actual locations. GCPs should preferably be 'hard' points: easily distinguished immobile feature edges such as buildings, houses, earth-bound anthropogenic structures, rather than 'soft' points such as vegetation, waterbodies, roads, or signs. A minimum of 5 to 8 hard GCPs is recommended (respectively, Mount et al., 2003 and Hughes et al., 2006) in order to minimize georeferencing uncertainty, with diminishing returns on uncertainty reduction beyond 8–10 GCPs (Hughes et al., 2006; Donovan et al., 2015; Lea and Legleiter, 2016). Homogeneous placement of GCPs reduces warping in georectified images, with moderate improvements when placing GCPs along the floodplain versus uplands (Lea and Legleiter, 2016). Multiple studies confirm that when georectifying images, a second-order polynomial interpolation minimizes error and reduces warping relative to higher-order transformations (Mount et al., 2003; Hughes et al., 2006; Lea and Legleiter, 2016). A summary of factors affecting georeferencing uncertainty is found in Table 1.

2.2.2. Approaches to river channel digitization and classification

Subsequent to quantifying uncertainty associated with georeferencing aerial imagery, the images are used to digitize the boundary of the stream via automated or manual methods. Regardless of the method used, channel-margin delineations are most often defined by the edge of riparian vegetation in order to eliminate variability arising from fluctuating water levels (Winterbottom, 2000; Gaeuman et al., 2003; Nelson et al., 2013; Rowland et al., 2016; Werbylo et al., 2017). The exception to this is when using delineations to estimate discharge or to determine if/how channel width has changed for a given flow value, in which case, variability as a function of water stage is desired (Bjerklie et al., 2005; Smith and Pavelsky, 2008; Lauer et al., 2017). Vegetated channel-margin delineations are more prone to error and inconsistency for braided and anastomosing systems, where width-related metrics are more sensitive to stage (Werbylo et al., 2017). An alternative approach to delineations uses the break in slope at the top of near vertical channel banks, but this approach depends on availability of high resolution topographic data and is unsuccessful where banks are not well defined (Osterkamp and Hedman, 1982; Mount et al., 2013; Donovan et al., 2015). While the edge of riparian vegetation provides a rational and typically discernable boundary, individuals must consider their data quality and research goals when defining an appropriate channel margin.

The scope of each project's questions and goals help inform whether to use manual or automated delineations, which can both help answer a wide variety of questions, but are each better suited for specific goals. Herein, we describe appropriate contexts and questions for each approach. Rowland et al. (2016) provided a detailed summary of methods and software using remotely sensed imagery to analyze river planform properties and dynamics (Micheli and Kirchner, 2002; Micheli et al., 2004; Güneralp and Rhoads, 2008; Legleiter and Kyriakidis, 2007; Aalto et al., 2008; Pavelsky and Smith, 2008; Lauer and Parker, 2008; Peixoto et al., 2009; Baki and Gan, 2012; Fisher et al., 2013; Hossain et al., 2013; Mount et al., 2013).

While manual delineations are most common (Blundell and Opitz, 2006) and accurate, semi- or fully-automated methods save time by eliminating the tedious nature of manual delineation (Güneralp et al., 2013; Rowland et al., 2016; Schwenk et al., 2017). Studies attempting to span broad spatial and/or temporal scales may therefore opt for automated delineations/classifications, thereby sacrificing accuracy to increase the extent of their analyses and save time. One caveat is that such studies must occur along sufficiently wide and active rivers (Peixoto et al., 2009; Constantine et al., 2014; Schwenk et al., 2017), because manual delineations are able to extend the spatial extent of delineations/classifications to lower-order streams relative to automated approaches. When conducted by an informed user, manual methods improve overall delineation/classification accuracy because they accommodate interpretation of occasionally complex or anomalous features and thus are flexible across varying hydrologic and geomorphic conditions. Because algorithms and automated

classifications adhere to a set of input criteria, they are more likely to misclassify channels or images outside the range of conditions for which they were developed.

Despite the limitations mentioned above, automated approaches such as SCREAM (Spatially Continuous Riverbank Erosion and Accretion Measurements; Rowland et al., 2016), are increasingly applied to multiple rivers with diverse morphologies and a broad range of image resolutions. SCREAM identifies and rasterizes channel locations from images to provide a suite of outputs including channel width and migration, sinuosity, bank aspect and channel islands (Rowland et al., 2016). In practice, automated channel classification algorithms such as RivWidth (Pavelsky and Smith, 2008) and SCREAM (Rowland et al., 2016), yield similar estimates of bankfull width for a variety of river planform morphologies. However, only SCREAM accounted for exposed channel bars and islands, and thus, width estimates had slightly higher discrepancies for multi-threaded channels.

Vectorized streambank delineations are often used to derive a single channel centerline. Changes in the location of the centerline over time can be used to estimate linear migration rates (Lauer and Parker, 2008; Donovan and Belmont, 2019; Sylvester et al., 2019). In other cases, polygons of the channel extent are delineated to measure areal change (Gaeuman et al., 2003; Rhoades et al., 2009; Donovan et al., 2015, 2016). Measuring migration as the difference in channel centerline position is a simple and efficient method to provide linear migration measurements over user-specified channel lengths. Linear migration rates and areal changes from polygons can both be normalized as a proportion of the channel width for comparisons across rivers of different size (Hooke, 1980; Donovan et al., 2015; Spiekermann et al., 2017; Sylvester et al., 2019). Estimating migration from centerlines conflates migration of each bank, and thereby reduces the ability to detect which mechanisms are influencing observed meander-bend migration (Miller and Friedman, 2009), as well as expected relations with flow (Schook et al., 2017). Thus, measuring migration separately for cutbanks on the outside of meander bends and point bars on the inside of bends is better suited for questions regarding local-scale mechanisms driving meander migration, at the expense of computation time.

An alternative to measuring linear channel adjustment is to compare polygons or rasters of channel position to estimate the area and/or volume of deposition and erosion (Gaeuman et al., 2005a, 2005b). This method is especially useful for braided or anastomosing systems, which may exhibit multiple linear adjustments within a single cross-section that should not be conflated. Channel polygons (vector or raster) are most often derived using image classification and assisted or unassisted machine learning algorithms. Such approaches benefit from the relative speed and ease of automation, but are limited by image resolution, shallow or transparent river reaches, inconsistent lighting and cloud cover, variable hydrologic or fluvial conditions, and vegetation and shadows overlapping the river boundary. Although raster and polygon outputs are generally similar, discrepancies increase between raster and vector-based delineations for lower order streams (Melville and Martz, 2004).

Differencing channel polygons is well-suited for measuring the area of erosion or accretion, but estimating volumetric sediment fluxes by combining linear migration rates with LiDAR or cross-section information consistently yields results similar to polygon-based calculations (Donovan et al., 2015; Rowland et al., 2016). Areas and volumes of erosion and/or deposition are most accurately and optimally estimated by differencing raster DEMs (DEMs of Difference; i.e., DoDs) derived from automated or semi-automated algorithms (Wheaton et al., 2010; Bangen et al., 2016; Vericat et al., 2017; Kelly and Belmont, 2018). However, such methods are limited to areas with repeat surveys of high-resolution topography (HRT), which are expensive to obtain and process due to equipment and software unavailable throughout the majority of history. Best practices and considerations for handling uncertainty in HRT DoDs are outside the scope of this paper, but can be found in Wheaton et al. (2010), Passalacqua et al. (2015) and Anderson (2018).

2.2.3. Uncertainty in river channel digitization and identification

A modest body of research has evaluated components of uncertainty for manual and automated channel boundary delineations and classifications (Downward et al., 1994; Gurnell et al., 1994; Melville and Martz, 2004; Rhoades et al., 2009; Liro, 2015). While automated channel delineations and classifications are increasingly common, manual delineations remain the most prevalent and accurate, and thus, serve as validation for estimating uncertainty of automated results (Rowland et al., 2016). Manual delineation error is greatly reduced by increasing digitization scale (i.e., zooming in), and to a minor degree, aerial photo scale (Liro, 2015). Mid-channel and side bars, as well as proximity to overhanging trees and shadows, are known to reduce the accuracy of semi-automated algorithms delineating wetted channel edges (Güneralp et al., 2013, 2014). However, the impact of such features and conditions on human delineation error is unknown; we attempt to fill this gap herein.

The magnitude, location, and type of offset due to inconsistent digitization and georeferencing distortions have unique implications for estimates of migration and width change (Rowland et al., 2016). For example, when bank delineations are biased in such a way that they result in a channel that is systematically narrower or wider than reality (Fig. 1a), centerline migration measurements will not be affected because the centerline(s) is/are not altered significantly. However, this scenario affects measurements of both channel width and width change. If the channel delineation is systematically offset in a single direction (Fig. 1a), channel width will remain unaffected, while migration will be exaggerated or damped. When only one bank is offset from its true location, measurements of width change will have an error equal to the magnitude of offset. However, the impact is halved for migration because the effect is damped when banklines are collapsed to a single centerline.

In the case of raster-based binary river classification, such as SCREAM, RivWidth or RivMap (Pavelsky and Smith, 2008; Rowland et al., 2016; Schwenk et al., 2017), 'delineation error' is equivalent to errors in feature identification and extraction. Each automation tool will have unique magnitudes and distributions of error due to unique classification algorithms and river width. For both automated classifications and manual delineations, the ability to detect migration will vary with image resolution, river size, rate of change, and time interval between photos. We hypothesize that georeferencing error will cause local systematic (i.e., directional) errors that impact migration measurements, but not necessarily width change measurements. However, over sufficiently broad reaches, georeferencing offsets should exhibit non-uniform/directional offsets. Due to the nature of changes in river width (i.e., narrowing and widening), width measurements are impacted by digitization inconsistency more than georeferencing offsets. We explain our approach to testing these hypotheses in the methods, and illustrate the results in subsequent sections.

2.2.4. Level of detection (LoD) threshold

The LoD is a threshold for determining whether measurements of migration or width change are statistically significant. This threshold should account for the total uncertainty arising from georeferencing (not orthorectification) and digitization uncertainty (or feature identification, in the case of automated classifications) (Mount et al., 2003; Mount and Louis, 2005). Research estimating migration and width change from repeat aerial images maintains a common practice of applying a uniform LoD threshold, typically ranging from 2 to 5 m based off the Root Mean Square Error (RMSE) or standard error (SE) of GCPs or delineations (Gurnell et al., 1994; Micheli and Kirchner, 2002; Mount and Louis, 2005; Rowland et al., 2016). Uniform error has also been estimated using photo scale (Gaeuman et al., 2003; Nelson et al., 2013), digitization scale (Downs et al., 2013; Liro, 2015), stated error from the source of the imagery (Draut et al., 2008), and qualitative

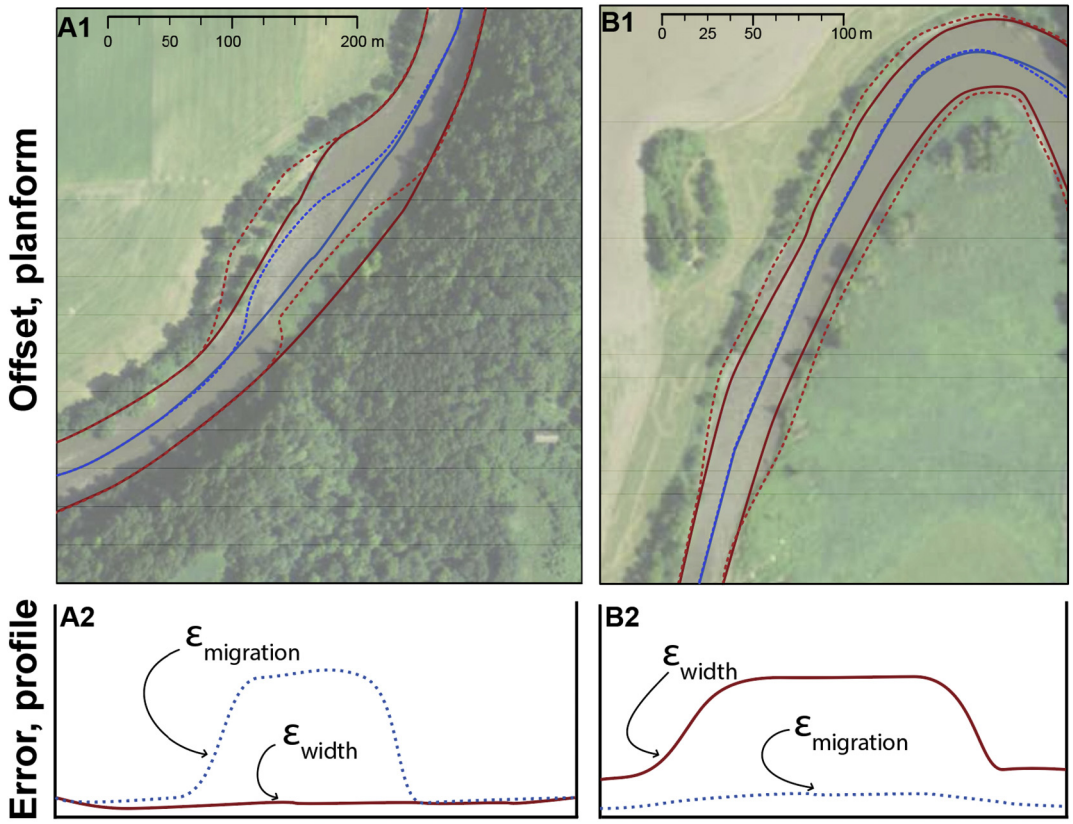

Fig. 1. Aerial imagery and hypothetical bank delineations illustrating the types of offsets (A1, B1), along with how such offsets have unique impacts on channel width and migration measurements (A2, B2). In Figures A1 and B1, the dashed lines indicate an alternative choice of delineation (red) and the centerlines associated with those banklines. The lines in Figures A2 and B2, illustrate- as a longitudinal profile- error in width (solid red line) and migration (dashed blue line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

descriptions of uncertainty in National Ocean Service T-sheets and U.S. Coastal Geodetic Surveys (Anders and Byrnes, 1991). However, uncertainty related to GCPs is inherently non-uniform and directional due to unique offsets of the GCPs in both x- and y-planes.

Three primary problems arise from assuming uniform LoDs, regardless of the how the LoD is estimated or what assumptions are made. First, actual at-a-point error is directionally skewed in the x- or y-plane (i.e., an ellipse), while uniform error thresholds can only be projected with equal error in all directions (i.e., a circle). Thus, uniform thresholds cannot accurately discern whether the error lies in the same or opposite direction of the channel migration or width change measurement. Second, RMSE-based LoDs increase loss of data/measurements relative to spatially variable LoDs (Liro, 2015; Lea and Legleiter, 2016), for reasons demonstrated in subsequent sections. Third, we hypothesize that uniform LoDs reduce the quality of retained measurements due to excessive loss of measurements of small changes and erroneous retention of large changes that may simply reflect georeferencing offset or image warping (Fig. 1b). These two issues are especially problematic for fluvial processes, which follow heavy-tailed distributions that are largely composed of small values and may be strongly skewed by a few large outliers. Discarding small but valid changes in combination with retaining large but erroneous changes causes a systematic overestimation of average migration rates.

Lea and Legleiter (2016) overcome the aforementioned issues by estimating spatially-variable (SV) uncertainty from both georeferencing uncertainty and digitization error. A raster of SV-uncertainty is used as an LoD to compare with migration trajectories from the Planform Statistics Toolbox (Lauer and Parker, 2008) in order to distinguish statistically significant measurements. Their approach calculates non-uniform LoD ellipses at all points around the river as the sum-of-squares including all sources of uncertainty (Eq. (1) and Eq. (2)).

$$\sqrt{\varepsilon_{xt1}^2 + \varepsilon_{xt2}^2 + \varepsilon_{dig}^2} = \varepsilon_x; \quad (1)$$

$$\sqrt{\varepsilon_{yt1}^2 + \varepsilon_{yt2}^2 + \varepsilon_{dig}^2} = \varepsilon_y; \quad (2)$$

$$\sqrt{\varepsilon_x^2 + \varepsilon_y^2} = \varepsilon_{xy}; \quad (3)$$

where ε_x and ε_y are georeferencing uncertainty in the x- and y-planes for time 1 ($t1$) and $t2$, respectively, ε_{dig} is digitization uncertainty, and ε_{xy} is total uncertainty at a location (e.g., the LoD error ellipse, Eq. (3)). This approach increased the number of statistically significant measurements retained relative to RMSE or 90th percentile uniform error thresholds (Lea and Legleiter, 2016).

Measurements below the LoD can be entirely removed, treated as values of zero (i.e., no migration occurred), or modelled based on the expected distribution of measured values (Donovan and Belmont, 2019). Replacing measurements below LoD with zero is generally preferable to removing them entirely. However, a disproportionate number (> 50%) of zero-values can compromise comparative statistics (e.g., Paired t-Test, Kolmogorov-Smirnov, Mann-Whitney Wilcoxon) that count each pair of zeros as a 'tie' (Martín-Fernández et al., 2003). In these cases, it may seem prudent to discard uncertain measurements, but this not only reduces sample size, but also introduces error into estimates of net change and uncertainty (Anderson, 2018). Details on imputation and replacement strategies for values falling below the LoD can be found in section 2.2.5.

Areal or volumetric estimates of channel change using polygons contain the same sources of error, but must account for error differently because outputs/results are 2-D geometric features (i.e., polygons, rather than 1-D lines). Polygonal results are commonly beset with small

polygon slivers that arise from slight misalignment due to georeferencing and digitization error (Chrisman, 1987; Bailey, 1988). Slivers are often assumed erroneous, rather than real change. Thus, a common and simple approach to address error is to remove polygons smaller than a threshold size (Edwards and Lowell, 1996; Donovan et al., 2015). More advanced approaches clip polygons using buffers of uncertainty scaled to georeferencing RMSE (Rhoades et al., 2009). Unfortunately, because both approaches lack a spatially variable threshold, both also unnecessarily remove significant measurements of erosion or deposition simply because they are of low magnitude. However, error models have been developed to incorporate heterogeneity of error for polygon/areal-based estimates of channel change (Gaeuman et al., 2005b). Rather than assuming errors are spatially independent of one another, they assume error is spatially correlated at the scale of local digitization inconsistencies and broadly over the scale of GCP placement, similar to Lea and Legleiter (2016).

2.2.5. Handling values below the LoD ('nondetects')

Earth-science literature increasingly reports uncertainty associated with planform change measurements, but there appears to be few discussions and no consensus regarding how to handle nondetect measurements that fall below the LoD/uncertainty threshold. Common methods include discarding nondetect measurements (Rhoades et al., 2009), imputing nondetects with values of '0' (Donovan and Belmont, 2019) or substituting nondetects with 0, 0.5, 0.7, or $1/\sqrt{2}$ of the LoD threshold (Martín-Fernández et al., 2003; Lee and Helsel, 2005; Helsel, 2006), and retaining nondetect measurements in order to bracket results with a range of upper and lower bounds based on the degree of uncertainty (Fraley et al., 2009; Dean and Schmidt, 2011; Donovan et al., 2016). Despite their use in contemporary research, these methods should be avoided because they introduce systematic errors and bias the mean and variance increasingly as the proportion of nondetects increases relative to the entire sample size (Tauber, 1999; Singh and Nocerino, 2002; Martín-Fernández et al., 2003). When observations or process understanding suggest no channel migration has occurred, it may be appropriate to record nondetect migration measurements as zeros, which are termed 'rounded zeros'. Researchers must consider, however, that the individual data points can no longer be log-transformed or used as a denominator in subsequent normalizations.

When nondetects account for only a small proportion (10–15%) of the data, EPA's Unified Guidance suggests that a simple substitution method is acceptable, based on insignificant changes to the mean and variance. However, when dealing with data fraught with nondetects, Maximum Likelihood Estimators (MLE), imputation via Regression on Order Statistics (ROS), and the Kaplan-Meier (KM) method are three approaches that provide more robust representations of summary statistics (i.e., mean, median, variance; Helsel, 2005). Each approach is unique, using some combination of detectable (i.e., known) data, nondetects (some methods, but not all), and an assumed or known data distribution based upon detectable observations. We describe the mechanics of each approach in the methods.

3. Methods & study area

3.1. Measuring migration and spatial autocorrelation

We develop guidelines to evaluate uncertainty for channel migration derived from manual channel delineations using a set of 13 aerial photographs spanning 76 years and 120 river-km of the Root River, MN, USA (Fig. 2A). With 441 images spanning nearly eight decades and an entire river network, we were able to evaluate the relationship between uncertainty (Eq. (1), Eq. (2), and Eq. (3)) and a variety of variables, including image resolution, acquisition date, local lighting, vegetation type/cover, and channel planform. Streambank delineations and interpolated centerlines from each georeferenced image (Souffront, 2014, M.S. Thesis) were used to calculate migration magnitude and rate at 10-m increments using the Planform Statistics Toolbox (Lauer, 2007; Lauer and Parker, 2008). The Planform Statistics Toolbox measures total migration as the distance between nodes on the initial and terminal channel centerlines, but does not identify meander bend cutoffs, which we manually identified and filtered out.

We initially measured migration along the entire reach at 10-m increments. However, measurements at such close intervals are likely autocorrelated due to the natural tendency for rivers to move in coherent spatial units that scale with the size of the river (Donovan and Belmont, 2019). Additionally, autocorrelation may arise from systematic offsets in digitization at local (100–500 m) scales, and at broader scales (10^2–10^3 m) due to offsets in image georeferencing and expected similarities in migration rates for adjacent stream reaches and meander bends. Autocorrelated measurements are not independent and are thus not statistically independent observations; they also underestimate standard errors and bias statistical comparisons that assume independent measurements. We computed Geary's C to estimate the length scale over which migration rates were influenced by the combined effects of spatial autocorrelation and local-scale systematic delineation biases (Geary, 1954). While Moran's I is more commonly used for characterizing global spatial autocorrelations, it fails to capture local autocorrelation due to the simple regression structure used. In contrast, Geary's C is able to detect local spatial autocorrelation, which is more relevant for analyzing channel planform adjustments, which are autocorrelated over meander-bend scale (local), rather than an entire longitudinal profile (global). Geary's C values typically range from 0 to 2 and estimate the level of correlation between all possible data points at specified lag distance bins. Values near to 1 indicate weak or absence of positive spatial autocorrelation. Values approaching 0 indicate positive autocorrelation, and are common at smaller lags. Values close to 2 indicate an increasing negative autocorrelation. To improve interpretability of results, we transform Geary's C values to standard correlation coefficients for plotting in correlograms, ranging from -1 to 1 to indicate negative to positive autocorrelation. Scripts for calculating Geary's C are provided in a supplementary file, 'GearyC.R'.

3.2. Georeferencing uncertainty

We analyze georeferencing uncertainty analyses using 13 sets of GCPs (n = 185–302) spanning 120 km of the Root River (Table 2). Rather than using the original GCPs, we quantified georeferencing error using an independent set of GCPs found on a high-resolution composite image from 2015 (USDA FSA APFO, 2015). Georectification transformations use least-squares fitting algorithms to optimize (i.e., minimize) georeferencing offset using the original input GCPs, but not necessarily areas in between (Ladd et al., 2006). Thus, using an independent set of GCPs ensured that error was not underestimated because it included areas aside from the original input GCPs. We primarily selected 'hard' GCPs (i.e., immobile or unlikely to have moved) that could be found on both historical and 2015 images. We evaluated spatial correlation of GCP error over a range of distances in order to determine whether or not georeferencing error is spatially correlated, and if so, over what distance (Gaeuman et al., 2005b).

3.3. Digitization uncertainty

We chose an 11-km reach of the Root River with multiple morphologic features and variable degrees of overhanging vegetation and shadows in order to quantify digitization uncertainty and determine whether it varies with fluvial and riparian conditions (Fig. 2B). The reach also allowed us to evaluate whether point bars reduced consistency in manual riverbank delineation, similar to semi-automated algorithms (Güneralp et al., 2013, 2014). Similar to (Gurnell et al., 1994), a single user repeatedly delineated 11-km of vegetation-streambank boundary four times without the aid of previous iterations, yielding 52 streambank delineations across 13 years with imagery

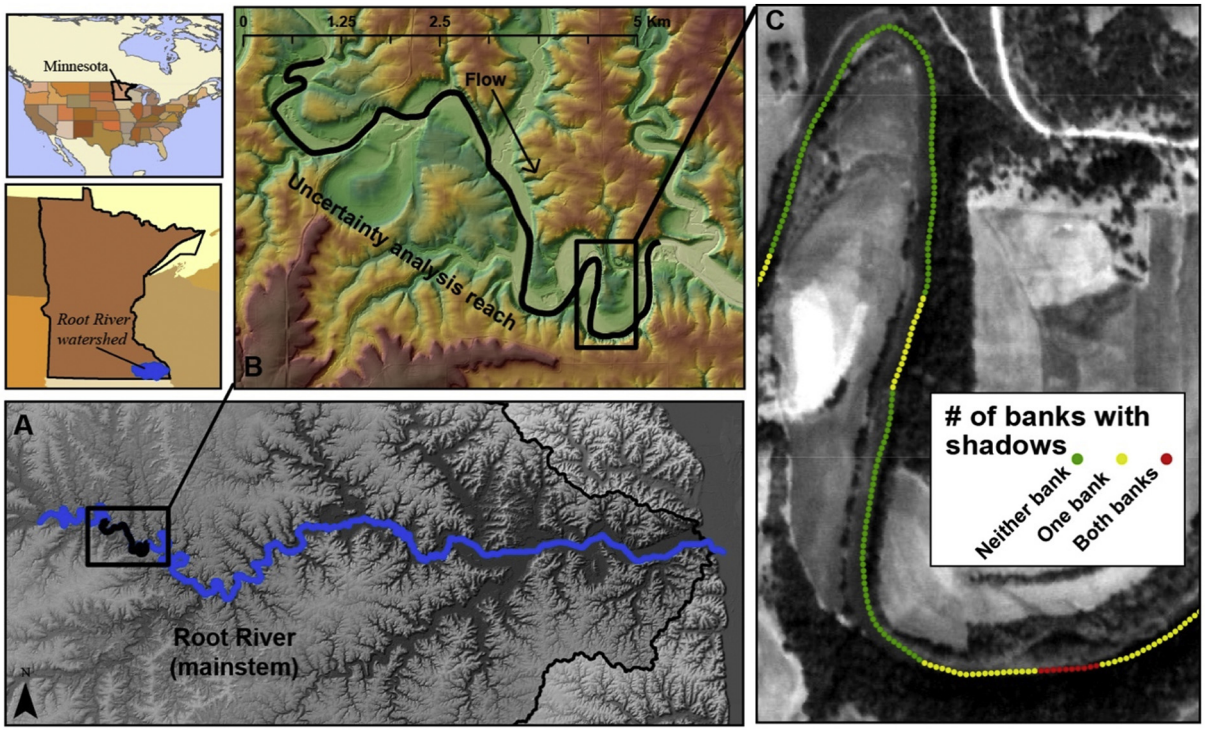

Fig. 2. (Topleft) Locations of Minnesota within North America, and the Root River within southeastern Minnesota. (A) 120 km of the mainstem Root River (blue) with 13 overlapping years of aerial photography. Inset black box is (B) the 11 km study reach analyzed in detail for this study. Second inset black box is (C) a single channel centerline divided into points each 10-m, with colors indicating whether shadows covered both, one, or neither streambanks. We categorized the banks in the same way for locations where vegetation covered the bank. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(13 years × 4 delineations). We did not limit the user to a specific map scale/zoom in order to reflect 'normal' working conditions, which do not impose such constraints. When vegetation covered the bank, we delineated bank location through the crown of the tree, unless we observed the bank-vegetation interface elsewhere. Channel centerlines were interpolated from each bank delineation, consistent with the prevalent method for measuring migration rates. Because centerlines should be identical for users with 100% delineation consistency, we calculated centerline offset at 10-m increments as a measure of uncertainty arising from digitization inconsistency. This resulted in 78 centerline comparisons (13 years × 6 centerline comparisons) spanning 11-km and 13 years of imagery, for approximately 86,000 measurements to estimate digitization uncertainty. We calculated Geary's C values from these centerline comparisons to quantify the length scales over which digitization introduces autocorrelation into migration measurements.

To determine whether image resolution influenced digitization uncertainty, first we visually compared the means and distributions of false migration of all 11-km of the 13 images using Kruskal-Wallis nonparametric analysis of variance (ANOVA) and a comparison of image resolution versus false migration. Subsequently, we evaluate whether a single user's digitization inconsistency increases for channel reaches with overhanging vegetation or shadows, as has been observed when using semi-automated delineation algorithms (Güneralp et al., 2013, 2014). At each 10-m increment, we classified the channel centerline as 0, 1, or 2; respectively classifying whether neither bank, one

Table 2
Overview of image characteristics, georeferencing control points, and error for each year with imagery. We used imagery from 2015 as a reference layer for calculating offset/error, and thus does not have error values. This assumes the 2015 imagery is the most spatially accurate representation of the study area.

Year	# of GCPs	Colour profile	Mean image resolution* (m)	Georeferencing error (m)				Digitization error* (m)		
				RMSE	10%	50%	90%	10%	50%	90%
1937	198	B&W	0.9	14.1	2.7	8.1	22.0	0.3	1.8	4.9
1947	214	B&W	1.0	9.3	1.9	6.0	14.5	0.2	1.4	3.7
1953	199	B&W	2.1	17.4	2.4	10.4	27.9	0.3	1.3	3.6
1976	185	B&W	2.2	18.8	4.5	13.0	31.4	0.4	2.2	4.9
1981	223	RGB	3.2	8.3	2.7	6.8	12.1	0.2	1.2	3.0
1991	220	B&W	1.0	3.8	0.8	2.9	5.9	0.2	0.9	2.9
2003	226	RGB	1.0	3.4	0.7	2.2	5.5	0.4	1.8	4.3
2006	227	RGB	2.0	4.5	1.2	3.4	7.0	0.4	1.6	4.7
2008	228	RGB	1.0	3.1	0.8	2.0	5.2	0.3	1.3	2.9
2010	232	RGB	1.0	3.5	1.2	2.8	5.2	0.3	1.7	4.4
2011	233	RGB	0.5	2.5	0.6	1.7	4.0	0.4	1.8	4.5
2013	233	RGB	1.0	2.9	0.9	2.2	4.6	0.3	1.4	3.2
2015	302	RGB	1.0	–	–	–	–	–	–	–

*Average resolution and digitization error come from the 11-km subreach.

bank, or both banks were obscured by shadow (Fig. 2C). We repeated this approach for vegetation, attributing 0, 1 or 2 to indicate absence or presence of vegetation obscuring one or both banks for each year with photo record. A Kruskal-Wallis test revealed whether any group/class exhibited significant differences. In cases where differences existed, we followed up with a Kolmogorov-Smirnov test to test for a stochastic increase in uncertainty for each class (Massey, 1951; Fay and Proschan, 2010).

3.4. Spatially-variable level of detection

We generated a spatially-variable level of detection (SV-LoD) raster for each year with images that included total uncertainty from georeferencing and digitization (Eq. (1), Eq. (2) and Eq. (3)). The sum of squares of positional offset in the x- and y-planes at each GCP and mean digitization inconsistency were interpolated using second-order polynomials, which yield low mean RMSE and minimize image warping relative to all eight possible transformations (Hughes et al., 2006; Lea and Legleiter, 2016).

We compared the percent and magnitudes of migration measurements retained (n = 66 comparisons) to respectively evaluate the relative quantity and quality of retained measurements. Comparing the percent of retention between SV- and uniform-LoD thresholds confirmed whether SV thresholds improved (i.e., increased) the quantity of retained measurements. We evaluated the quality of retained measurements by testing whether distributions under SV thresholds shifted left (i.e., reduction) relative to uniform LoD thresholds. This assumes that a reduction in the distribution indicates that the quality of retained measurements has increased because SV-LoDs are more likely to retain measurements of small magnitude that are erroneously discarded by uniform LoDs. We visually inspected shifts in addition to performing one-way Kolmogorov-Smirnov tests to evaluate whether SV distributions increased the retention of low-magnitude measurements relative to uniform LoD thresholds, as indicated by a reduction/leftward-shift in the distribution. Lastly, we used linear regressions to estimate percent retention as a function of pixel resolution, initial image year, temporal measurement interval, along with river-averaged migration and its natural logarithm.

3.5. Handling nondetects

Common approaches for handling nondetects in geomorphology include: removing nondetects or substituting nondetects with 0, 0.5, 0.7, or $1/\sqrt{2}$ of the LoD threshold (Martín-Fernández et al., 2003; Lee and Helsel, 2005; Helsel, 2006) and retaining nondetect measurements in order to bracket results with a range of upper and lower bounds based on the degree of uncertainty (Fraley et al., 2009; Dean and Schmidt, 2011; Donovan et al., 2016). However, these methods introduce error into estimates of net change and uncertainty (Anderson, 2018) and skew statistical parameters by introducing a disproportionate number of arbitrary values (Helsel, 2006), respectively. Thus, we propose and test three alternative methods developed by statisticians and discipines outside of Earth science: Maximum Likelihood Estimation (MLE), imputation via Regression on Order Statistics (ROS), and Kaplan-Meier (KM). Each method is unique in its approach to estimating summary statistics for nondetect data and are known to outperform one another depending on the underlying data distribution, proportion of nondetects, sample size, and the number of detection limits (Helsel, 2005). Among research on geomorphic change detection, bracketing results with the sum of uncertainties is common; the benefits and drawbacks of this approach are detailed in (Anderson, 2018). Our scripts for running MLE, ROS and KM are available in the supplementary. R file.

Following its name, the MLE estimates the 'most likely' mean and standard deviation by fitting both detected observations and nondetects to a distribution chosen by a knowledgeable expert. Although MLE assumes a normal distribution, it is commonly used with transformed lognormal data. MLE generally underperforms for small datasets (n < 50 detectable observations) with large skewness, or when outliers are present, relative to ROS or K-M. ROS depends less on assumptions of distribution shape because it estimates nondetect data using probability plots of detectable data. Kaplan-Meier is a standard in medical, industrial, and water chemistry statistics for estimating the mean and standard deviation of data containing censored (i.e., partially known) measurements. K-M does not assume a parametric distribution, but requires at least 8–10 measurements, < 50–70% nondetects, and is biased when the highest and/or lowest values are nondetects. K-M also requires multiple levels of detection, and thus, is appropriate with an SV-LoD, but not a uniform LoD. For a robust description of K-M, see Hosmer et al. (2008). Additional guidance and details on MLE, ROS, and K-M are provided online (Huston and Juarez-Colunga, 2009; ITRC, 2013).

We evaluated each approach by quantitatively comparing their predicted mean (μ), median, distribution fit, and standard deviation (σ) with known values from modelled distributions (n = 400) containing varying proportions of nondetects (8–30%). The 'best estimate' of mean, median and variance from the MLE, KM, and ROS were those with the minimum difference relative to the modelled/raw values. We ranked the distribution fits relative to the original modelled distribution using a Pairwise Wilcoxon Rank Sum test statistic with an adjusted p-value (based on Benjamini and Hochberg (1995)) to reduce the rate of false-positive results and allow for distribution comparison. With this adjustment, the test is sensitive to differences in distributions, not only central tendencies. We plot the empirical cumulative density functions (ECDFs) from each approach to visually confirm the quantitative results and further inform a discussion of when each approach is most appropriate. We also varied the measurement sample size (n = 100, 1000, or 10,000) in modelled simulations to evaluate whether sample size influenced which approach (i.e., MLE, KM, ROS) best predicted the statistical parameters. This also allowed us to explore the implications of different combinations of spatial and temporal data extent (spatial coverage x measurement interval), which directly affect sample size.

Modelled migration rates followed a lognormal distribution with means drawn from the range of 13 empirical distributions, each with ~13,000 measurements spanning 120 km of the Root River. Deviance scaled directly and significantly with the mean migration rates, and thus was predicted using Eq. (4), similar to Donovan and Belmont (2019).

$$\sigma = 0.25\mu + 1.09 \quad (4)$$

We calculated the probability of significance for each modelled migration rate using the relationship ($r^2 = 0.89$) between empirical migration rate and chance of statistical significance based on 864,204 empirical migration rates (Fig. 3). We did not include migration rates beyond 9 m/yr when regressing the data because 100% of those rates were significant, and including those reduced the logarithmic regression fit for values that may be nondetects. This approach resulted in 10 to 53% of nondetects for each model iteration, which had 100, 1000 or 10,000 sample measurements.

4. Results & discussion

4.1. Spatial autocorrelation for measurements of migration and uncertainty

Correlograms of autocorrelation values illustrate a waning spatial autocorrelation of channel migration rates over length-scales that are approximately 1–4 channel widths (50–200 m), at which point autocorrelation was weak or nonexistent (Fig. 4). The trends in autocorrelation values of migration rates were similar to those for user digitization inconsistency. Thus, we were unable to ascertain the length scales over which autocorrelation reflects digitization inconsistency as opposed to coherent units of channel migration. This suggests similar

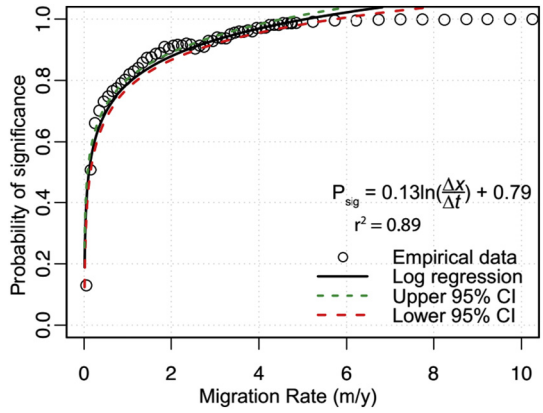

Fig. 3. Empirical data (black dots) used to generate the probability that a modelled migration rate will be significant or nondetect. The probability of significance (P_{sig}) for a given migration rate increases logarithmically with the magnitude of the rate, and beyond ~9 m/yr, 100% of measurements were significant. In order to improve the model fit for values with a chance of being nondetects, values with 100% chance of significance were not included in the regression.

scales of autocorrelation for both manual channel digitization and river migration across a wide range of geomorphic conditions and river widths. Thus, measuring migration over lengths ≥6 channel widths (400 m) for the Root River ensured autocorrelation did not confound statistical results and inferences.

When testing the assumption that GCPs exhibit local spatial correlation (Gaeuman et al., 2005b), we found strong autocorrelations were rare (values > 0.7 or < −0.7) (6–28%), whereas weak autocorrelations dominated the data (−0.3 > r < 0.3) (62–81%), with the remaining 14–38% of data exhibiting moderate autocorrelations (values between 0.3 and 0.7 or − 0.3to − 0.7). The few GCPs with strong autocorrelation did not dominate any particular spatial scale, suggesting it is invalid to assume that nearby GCPs are more or less similar than distant GCPs. The lack of local or global autocorrelation reinforces the need for spatially variable LoDs because neither neighboring GCPs nor distant GCPs were similar in magnitude or autocorrelated.

4.2. Factors influencing digitization uncertainty

We analyze an 11-km reach of the Root River to build upon the literature quantifying the factors that influence the magnitude, location, and types of offset from digitization. Mean and median digitization uncertainty across all years was 1.4 m. Despite unique approaches to calculating uncertainty, this result is consistent, albeit slightly less than Gurnell et al. (1994) (2 m). Within the 11-km segment, pixel resolution ranged from 0.5–5.8 m^2, thereby including nearly the full pixel range (0.3–5.8 m^2) for all 441 images spanning the 120-km mainstem Root River. Despite a variety of pixel resolutions, the distributions of digitization offset were not significantly different (p = 0.95, Fig. 5) based on a Kruskal-Wallis Rank Sum Test and the lack of a systematic trend between digitization offset and pixel resolution. Neither image date nor pixel resolution appear to have a systematic influence on the degree of digitization inconsistency for a single user. The consistency in these results across a variety of image resolutions (0.5–5.8 m^2) and conditions suggests that average and median digitization uncertainty for a single experienced user will lie between 1.5 and 2 m, and need not be calculated for all future studies. It remains plausible that follow-up evaluations may be pertinent for studies with resolutions outside this range, or for substantially different geomorphic conditions. The framework provided herein for evaluating digitization uncertain remains transferrable to other environments to explore potential differences. Similar methods were applied to braided and anastomosing planforms by Werbylo et al. (2017), who found that measurements of at-a-section channel width derived from multiple digitizers differ up to 20% of channel width, while river-averaged widths exhibited no significant differences.

Digitization uncertainty was generally consistent across all years (i.e., high precision) in cases where the bank is masked by shadow and/or vegetation cover (Fig. 6). Higher degrees of inconsistency (> 5 m) occurred along meander bends with various types of vegetation cover (e.g. thick vegetation, scattered brush, grass) because users are inconsistent in their choice of vegetation boundary. Similar inconsistencies have been observed for semi-automated algorithms (Güneralp et al., 2013, 2014). Most of the remaining offsets in digitization (≤ 1 m) were minor and were scattered uniformly across the 11-km reach. Thus, we demonstrate that users remain consistent in delineating the channel-vegetation boundary regardless of shadows and vegetation, but not in cases where multiple vegetation boundaries exist. Thus, for such reaches, users should determine which vegetation-boundary best reflects the dominant- or channel-forming discharge. Optimally, verification would utilize ground-truthing along ambiguous reaches. Alternatively, where high-resolution topography is available, local peaks in curvature may be used as a characteristic signal of the streambank-floodplain transition (Donovan et al., 2015). In areas where riparian vegetation and geomorphic conditions differ substantially to obscure delineations, the framework laid out herein should be applied to evaluate for consistency of results.

4.3. Georeferencing uncertainty

Georeferencing uncertainty varied widely across the study reach within each year's images (Fig. 7) and across all years (Fig. 8), and generally decreased towards the present, likely a reflection of reduced distortion and warping with improved camera lenses and developments in self-calibrating sensors (Clarke and Fryer, 1998). Because

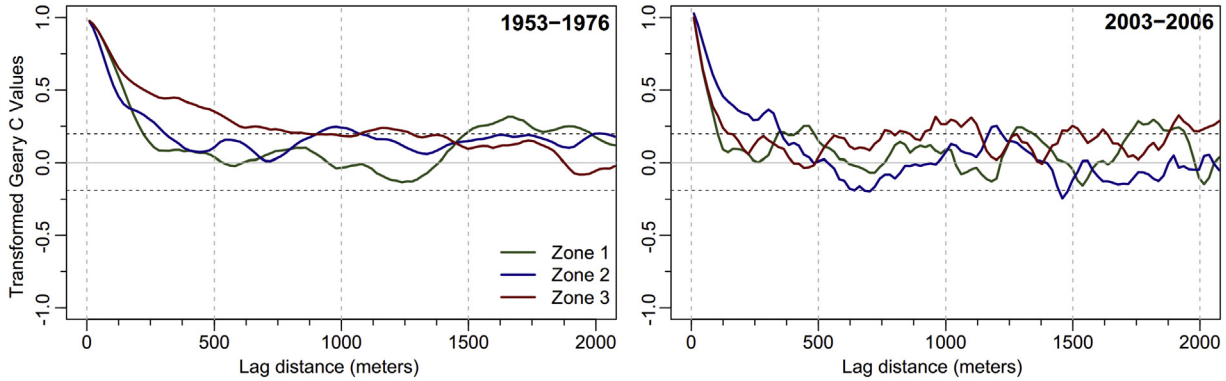

Fig. 4. Correlograms of spatial autocorrelation for measurements of channel migration. Lag distance indicates the length over which autocorrelation was measured. Geary's C values, which typically range from 0 to 2, were transformed to the typical range of correlation values, spanning −1 to 1.

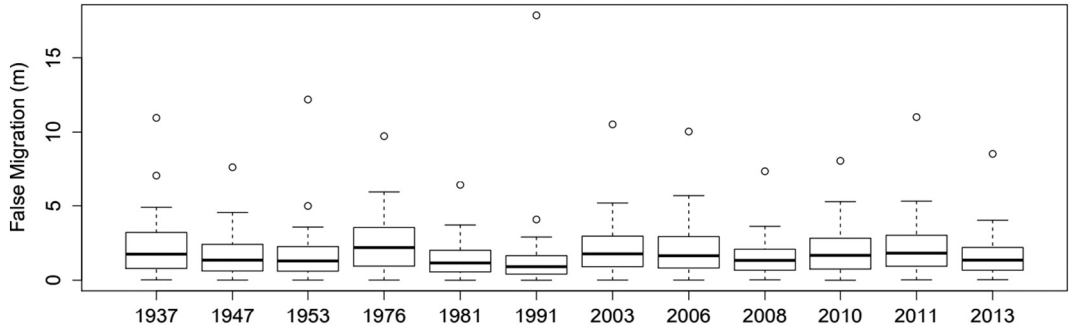

Fig. 5. Digitization inconsistency for 11.2 km of the Root River for each of the 13 years with aerial images (~1120 migration transects measured for each photo year). The overall magnitudes and distributions were relatively consistent across all photos, regardless of pixel resolution.

georeferencing error generally distorts images over scales equal to and greater than GCP spacing, river width measurements are unaffected unless warping occurs over scales less than a channel width. Thus, we do not explore the impact of georeferencing uncertainty on width calculations. Neither the RMSE of georeferencing error, nor the distributions of error, appear to have any significant relationship to the mean pixel resolution of each year (Table 2) based on graphical assessment, and thus, were not explored statistically. The RMSE was approximately the 75th percentile for most years (Fig. 8, red points), which illustrates how a small number of extreme outliers in georeferencing uncertainty inflate the RMSE relative to the median (i.e., the actual central tendency in a long-tailed, non-normal distribution). We describe implications of using this inflated RMSE value in the next section.

4.4. Calculating and evaluating LoD thresholds

Final LoD thresholds reflect the sum of squares of georeferencing and digitization uncertainties. While SV LoDs reflect a different sum of squares for each pixel, RMSE LoDs calculate a single, average sum of squares. As explained earlier, however, extreme outliers in georeferencing error disproportionately inflate the mean error (i.e., RMSE) relative to the majority of values within a lognormal distribution (Fig. 8, red dots). Thus, using an inflated RMSE value as the LoD inherently removes the majority of migration measurements, which are predominately low-magnitude values (Fig. 9, black distribution). After using a singular RMSE LoD, the majority of low-magnitude migration values are thus 'nondetects', while the few upper percentiles remain to dominate the migration distribution. As a result, the mean migration rate- a widely used metric for comparing rivers, estimating streambank sediment loads, and constraining sediment budgets- is overestimated.

Summary statistics for each year's imagery and the associated georeferencing and digitization uncertainty are available in Table 2. Second-order polynomial interpolation of SV error improved the number of retained migration measurements for all years,

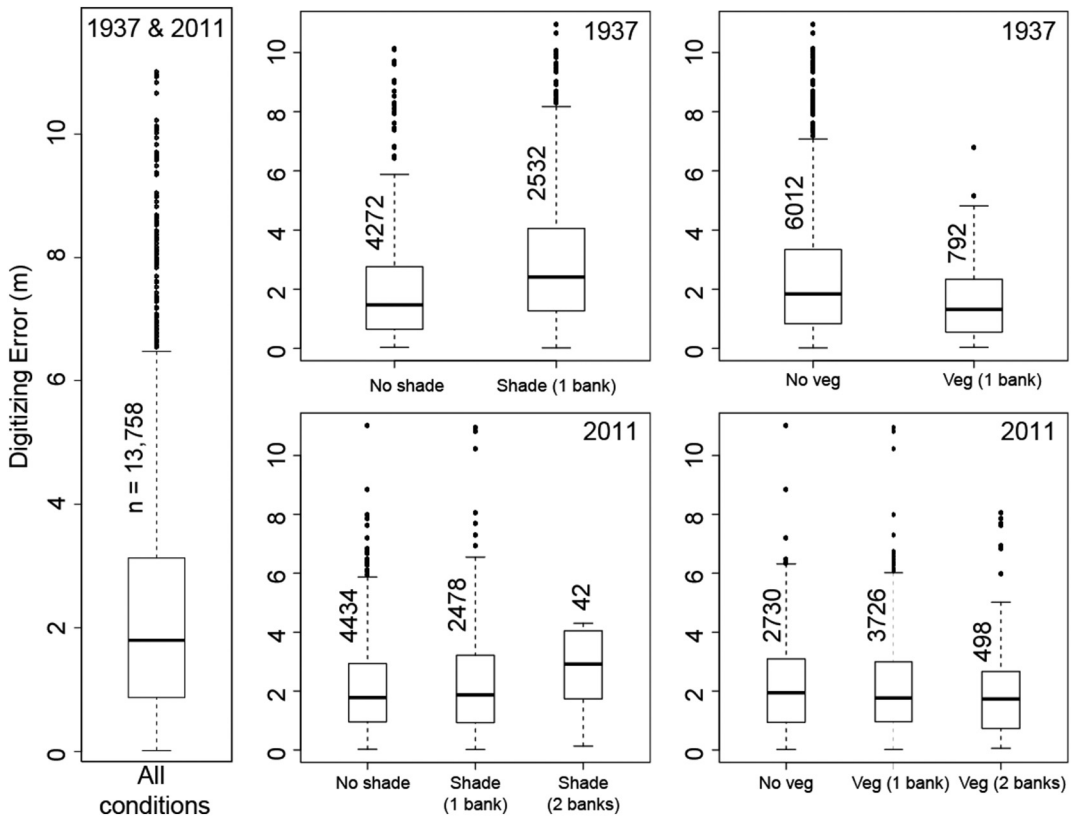

Fig. 6. Digitization inconsistency binned by the presence of shadows and vegetation on either one, both, or neither adjacent streambanks. Numbers adjacent to individual boxplots indicate the sample size of each category.

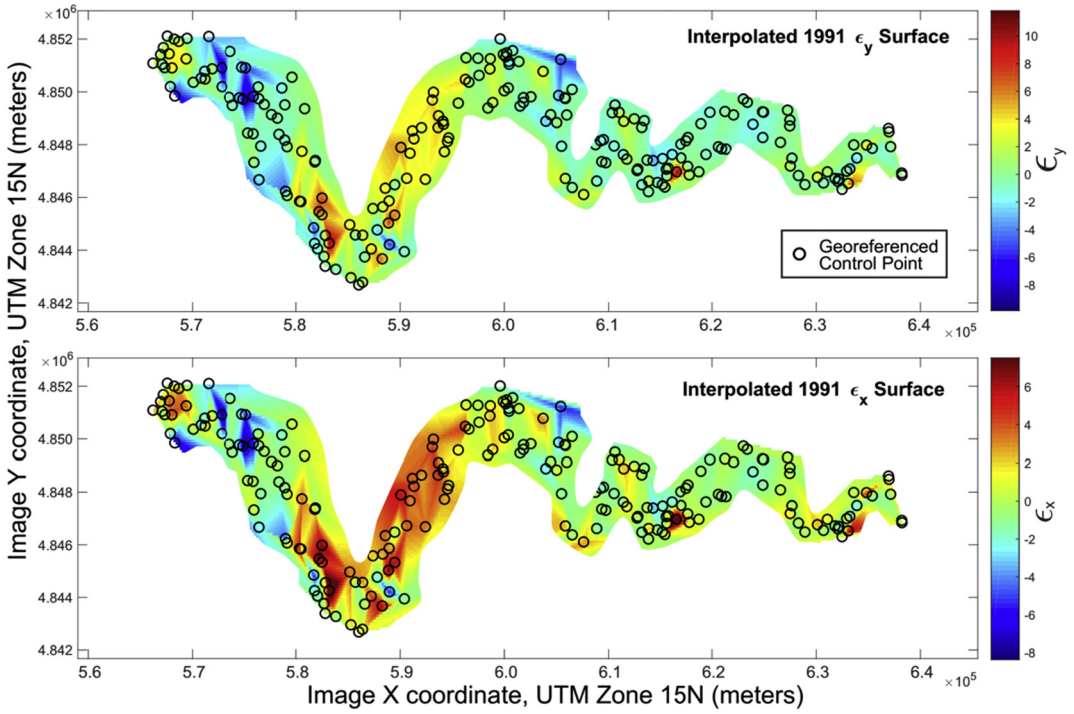

Fig. 7. Spatially variable georeferencing uncertainty, in meters, across the x- and y-coordinate planes (bottom and top, respectively). For total uncertainty, we used Eq. (1) and Eq. (2) as components for calculating a final error ellipse for each pixel. Note that the colour gradient scales change for each panel, due to different ranges of error in the x- and y-planes.

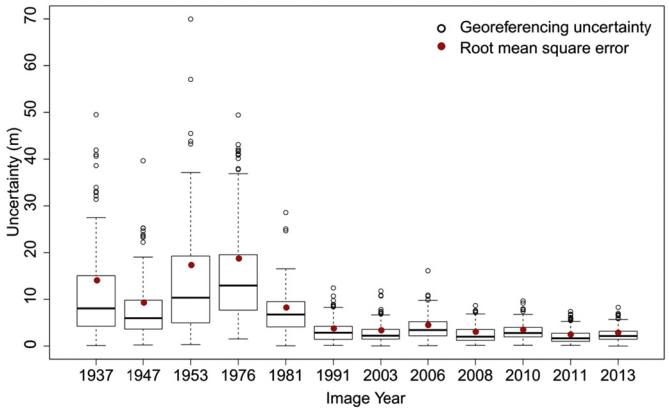

Fig. 8. The distribution of georeferencing uncertainty for each image year (black), based on the set of georeferenced control points (GCPs, n = 185–302). The red dots indicate the root mean square error (i.e., mean) of each distribution, a common uniform uncertainty threshold. However, using the RMSE threshold leads to excessive data loss of any measurement below the red dots. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

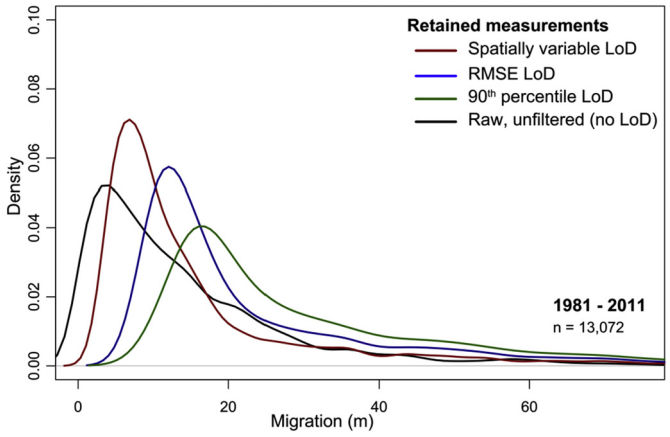

Fig. 9. A comparison of probability density functions among of all measurements (black) and retained (i.e., 'significant') measurements for each type of LoD (SVE-red, RMSE-blue, and 90th percentile threshold-green). Illustrated distributions are for measurements of migration (m) between 1981 and 2011, with similar trends in the remaining 65 comparisons. The SV-LoD (red) retained a higher proportion of low-magnitude measurements, and relatively fewer large-magnitude measurements. This suggests that a SV-LoD not only improved the quantity, but also the quality of retained measurements. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

corroborating results from the original implementation of this method (Lea and Legleiter, 2016). Furthermore, in all comparisons (n = 67), SV-LoD thresholds retained more migration measurements (μ = 62%, range = 25–81%) than uniform LoD thresholds (μ = 35%, range = 12–52%). Distributions of retained measurements consistently demonstrate that SV-LoDs retain additional measurements of smaller magnitude and fewer measurements of large magnitude (Fig. 9).

For three different years, we evaluated migration measurements that were retained with the SV-LoD, but not the uniform LoD. Visual observations of values retained by only the SV-LoD suggest that a little over half of these values were real, verifiable changes, generally characterized by gradual systematic shifts in the river apparent in multiple images and/or visual evidence from the LiDAR hillshade and/or images. 'Questionable' retained measurements were often non-systematic migration along reaches that were fully or partially masked by shadows or vegetation, which possibly led to inaccurate delineations. Thus, despite consistent delineations (i.e., delineation precision) under the same conditions, the interpreted location of the riverbanks (i.e., delineation accuracy) still appears to vary significantly when visual conditions change along a stable reach (Fig. 3 and Fig. 6).

Anomalously high-magnitude measurements that result from extreme image warping or georeferencing error that shifted the river

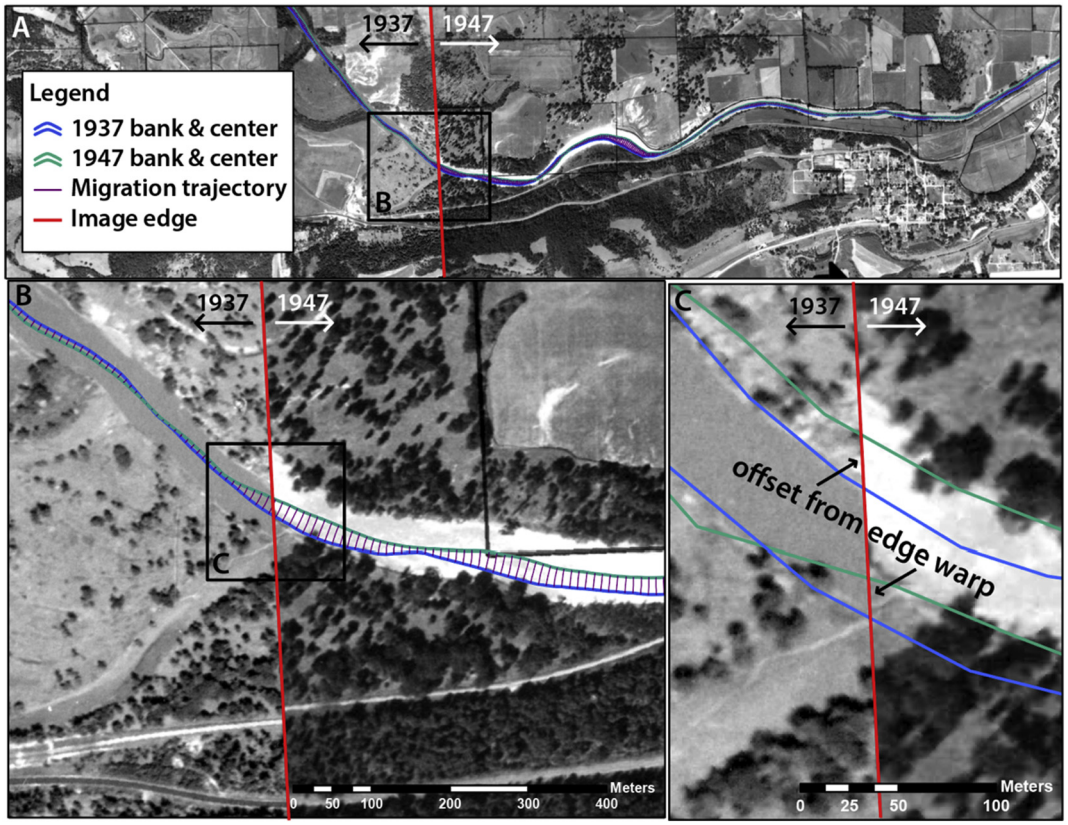

Fig. 10. An example of image warping near the edge of two images, which intersect at the red line. The warping resulted in offset channel centerlines (B) and delineated banklines (C). Evidence of warping along the channel in B and C is confirmed by the offset roads in the bottom of B. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

boundary (Fig. 10) are more likely to be retained by the SV-LoD than the uniform LoD. The uniform LoD may retain these large measurements simply because they exceeded the RMSE, whereas the SV-LoD is likely to discard them by properly accounting for local maxima in uncertainty. The use of a spatially-variable LoD will thus improve results by reducing the likelihood of including inaccurate measurements, and is generalizable/applicable to any context.

We tested for systematic trends in the percent of retained measurements to evaluate whether factors such as image year, image resolution, or migration distance, exhibited strong relationships with the proportion of retained measurements. Of these factors, the most significant was the natural logarithm of mean migration distance (Fig. 11, $p < 0.001$, $r^2 = 0.92$). When mean migration falls below the minimum LoD- equivalent to mean and median digitization uncertainty (1.4 m, dashed-blue line, Fig. 11)- there is a drastic reduction in the proportion of 'significant' migration measurements. For migration above this threshold, the number of significant measurements increases relatively slowly, as some diminishing proportion of measurements still have a total uncertainty that exceeds the LoD threshold.

4.5. Treatment of nondetect measurements

We summarize the model results for each approach- Maximum Likelihood Estimator (MLE), Kaplan-Meier (KM), Regression on Order Statistics (ROS), zero thresholding, and removing nondetects- to best approximate the mean, median, standard deviation, and overall distribution of modelled data in Table 3. We quantified 'best fits' for the mean, median and standard deviation as the method which best approximated (i.e, minimum difference) a known, modelled distribution. We ran 400 simulations, each with a unique lognormal distribution

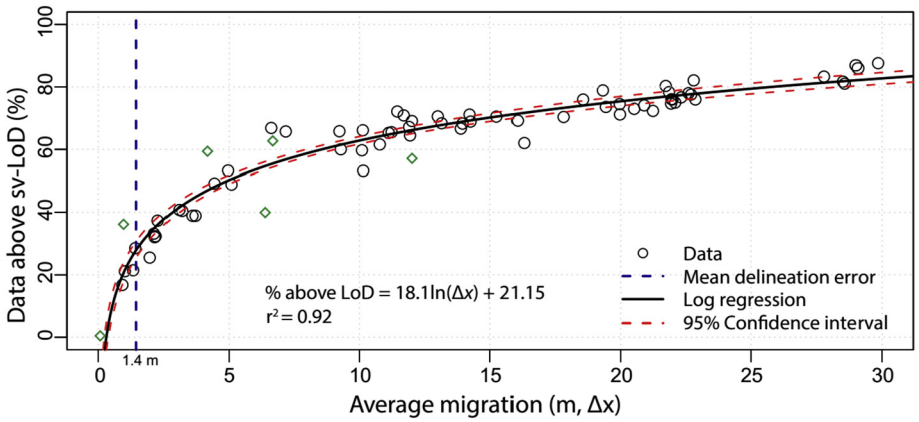

Fig. 11. Relation between mean migration and the percent of retained measurements. Each black point reflects the percent of significant measurements for a given river-averaged migration (over the 120-km of the Root River) between two sets/dates of imagery (n = 864,204). The proportion of significant measurements increases rapidly as mean migration passes the average (mean and median) digitization error (blue vertical line, 1.4 m). Beyond this threshold, the number of retained measurements increases asymptotically. Green diamonds represent averaged migration measured over portions (1–37 km) of the Root River, rather than the entire 120-km (black points). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Results from 400 simulations comparing known/modelled values with estimated statistical parameters (i.e., mean, median, standard deviation, distribution) from each method. MLE- Maximum likelihood estimator; KM- Kaplan Meier; ROS- Regression on Order Statistics; ZRO- nondetects treated as '0'; RMV- nondetects removed entirely. Bolded values indicate which method performed the best.

Sample size	Statistical criteria	Best fits (%)					
		MLE	KM	ROS	ZRO	RMV	Total
n = 100	Mean	**50%**	33%	0%	0%	17%	100%
	Median	22%	**41%**	1%	2%	34%	100%
	Variance (σ)	2%	7%	**80%**	0%	11%	100%
	Distribution	**91%**	1%	0%	8%	0%	100%
n = 1000	Mean	**50%**	34%	1%	0%	15%	100%
	Median	1%	**77%**	0%	0%	22%	100%
	Variance (σ)	1%	1%	**98%**	0%	0%	100%
	Distribution	**99%**	0%	0%	1%	0%	100%
n = 10,000	Mean	**47%**	30%	1%	0%	22%	100%
	Median	0%	**79%**	0%	0%	21%	100%
	Variance (σ)	0%	16%	**84%**	0%	0%	100%
	Distribution	**98%**	0%	0%	2%	0%	100%

containing a range of nondetects (47–90%) and sample sizes (n = 100, 1000 or 10,000). Detailed methods for generating and comparing the statistical parameters are described in Section 3.5, and can be found in the supplementary. R file "Nondetect Method Comparison'.

Across all sample sizes, the MLE and KM methods provide the best estimate of the mean in a similar number of simulations. MLE performed better for model iterations with higher proportions (20–40%) of nondetects, whereas KM performed best with lower proportions (10–20%) of nondetects. This is intuitive, due to the way MLE handles nondetects in estimating the new mean. KM also performed well in estimating the median across all sample sizes with the exception of small sample sizes containing a high proportion of nondetects, where removing nondetects (RMV) resulted in better estimates of the median. In the majority of simulations (86–97%), ROS provided the best approximation of variance. Finally, the best estimates of the entire distribution largely came from the MLE (85–92%). Two characteristic sets of ECDFs are plotted alongside the raw/modelled data as the gold standard (Fig. 12). Below, we illustrate and describe the conditions in which each method is most appropriate in a flow chart (Fig. 13).

5. Conclusions, recommendations, and future challenges

Earth-science literature has become increasingly aware of the importance of calculating and disclosing uncertainty inherent in GIS-based measurements. In the early 1990s, (Anders and Byrnes, 1991) acknowledged the need to address the key sources of uncertainty in boundary delineations derived from aerial images. This body of research grew throughout the following decade to describe how to estimate uncertainty and levels of detection (LoDs) using traditional methods of error propagation (Edwards and Lowell, 1996; Kiiveri, 1997; Crosetto and Tarantola, 2001). Subsequent research quantified how specific variables influence uncertainty (Gaeuman et al., 2003; Nelson et al., 2013; Güneralp et al., 2013, 2014; Liro, 2015). Meanwhile, other fields and researchers were identifying appropriate methods for handling nondetect measurements below LoD thresholds (Shumway et al., 2002; Martín-Fernández et al., 2003; Helsel, 2006; Lee and Helsel, 2005). Unfortunately, a disconnect remains between these developments and the application within earth-science research (Lea and Legleiter, 2016; Donovan and Belmont, 2019). Herein, we have collected, summarized, and tested methodological and applied research relevant to calculations of planform changes derived from remotely-sensed imagery. The subsequent paragraphs provide a comprehensive framework including both general guidance and specific factors to consider when evaluating uncertainty in planform change measurements.

Our use of 441 images spanning 8 decades and encompassing multiple riparian conditions, geomorphic environments, and a wide array of resolutions and image quality (e.g., grayscale and colour) allow us to provide a general framework for handling uncertainty that is broadly applicable to rivers of varying scale, geomorphology, and river pattern. Nonetheless, we recognize the need to include considerations and caveats for our specific analyses, which stem from a single threaded meandering river (Root River, MN) spanning widths of 30–80 m. Many of these considerations are detailed in the background material, but are restated in the context of our results, below. While specific unforeseen considerations will vary with each application, the practices, conclusions, and recommendations for calculating and evaluating uncertainty are generally applicable to many remote sensing applications, including glacier retreat or advance, erosion or deposition along coastlines and lakeshores, changes in wetland extent, expansion or contraction of vegetation (e.g., deforestation), cliff retreat, and political boundary disputes. We encourage readers to consider their specific context, questions, and needs when applying our findings. To reconstruct the analyses conducted herein to explore their own datasets, we have made scripts are available as supplementary. R files. For example, our approach focuses on measuring change as a linear adjustment, but as explained in the background, anabranching and braided channels exhibit complex planforms and adjustments may be better approximated by a volume, mass, or percent of erosion/deposition.

Surprisingly, inconsistencies in streambank delineations were generally not significantly greater for images of lower quality (resolution) or channel reaches with shadows and/or vegetation cover. Thus, a single, experienced user can be expected to have a similar degree of precision (i.e., consistency) regardless of image quality, shadows, and/or vegetation, at least for settings with a range of environments similar to those encountered along our 120 km stretch of the Root River. Arbitrary inconsistency in user-defined delineations dominate delineation uncertainty, but we expect that image quality will dominate

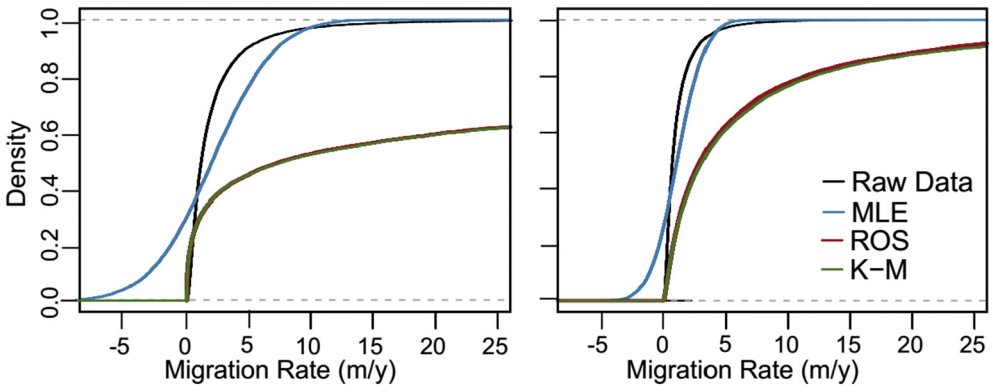

Fig. 12. Empirical cumulative density functions (ECDFs) for the modelled (i.e., 'raw') migration data, alongside three approaches used to model nondetect measurements. The two sets of ECDFs reflect the majority of model iterations (n = 400). While the three methods are only meant to estimate summary statistics (mean, median, variance), visualizing the distributions help interpret the results in Table 2.

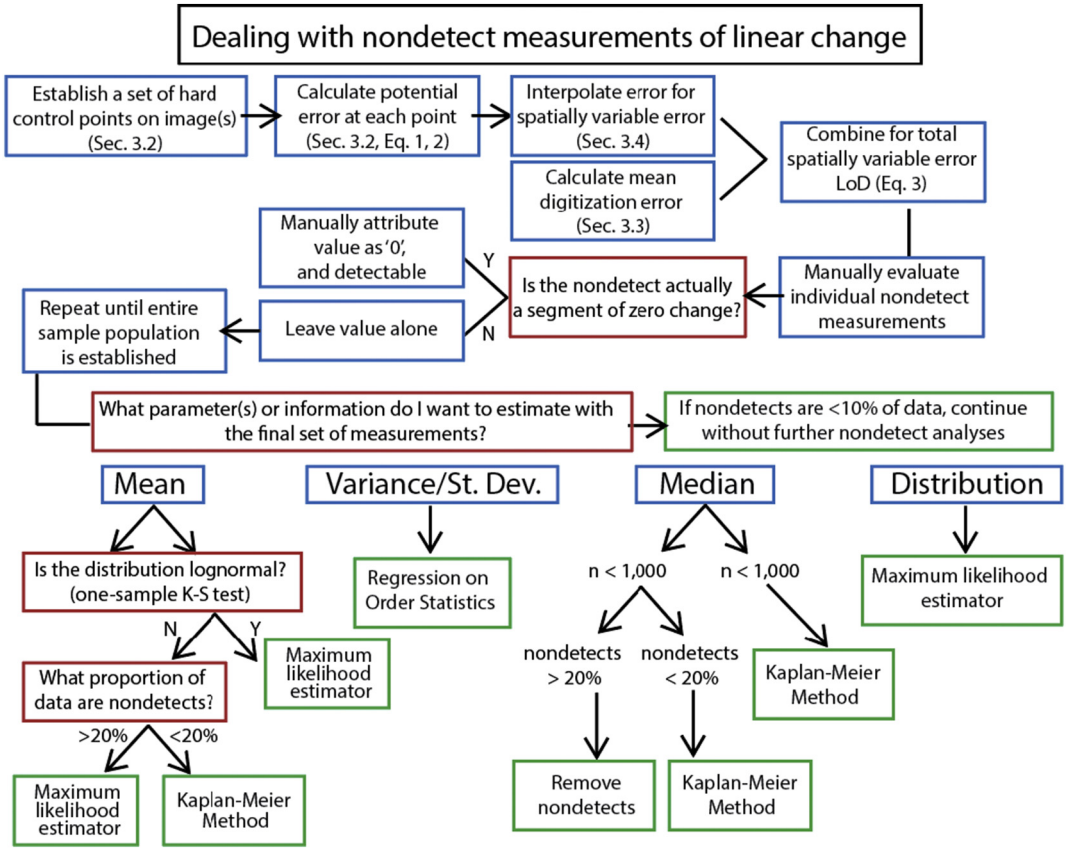

Fig. 13. A flow chart illustrating how to handle nondetect measurements, from start to finish.

delineation uncertainty when pixel size exceeds the resolution necessary for detecting the riparian-fluvial boundary (0.5–3.5 m^2 were used herein), such as in (Werbylo et al., 2017). Vegetation type, or an absence thereof, will also impact whether pixel resolution leads to more or less accurate delineations. Furthermore, pixel size may exceed or span the width of small tributaries and/or sufficiently narrow channels, which exacerbate the impact of small delineation offsets per channel width.

In cases where image quality weakens or invalidates delineations, field measurements or high-resolution topography should be used to verify or replace image-based river delineations. Locations along meander bends with various types of vegetation cover (e.g. thick vegetation, scattered brush, grass) exhibit the highest uncertainty due to ambiguity in which vegetated boundary to use. Future studies should explore how to incorporate spatially variable delineation uncertainty. Performing digitizations with a single user is optimal because it reduces digitization uncertainty by approximately 0.5 m relative to multiple users, which has been shown to have a central tendency of ± 2 m (Micheli and Kirchner, 2002) and range up to ± 37 m (Werbylo et al., 2017). Specifying a standard for delineation will help reduce errors and biases in long-term monitoring of river channels and riparian conditions that rely on multiple users for manual delineations. We recommend delineating the vegetated boundary that best approximates bankfull width to avoid inconsistency along such reaches, whenever possible. Additional considerations may be necessary for delineations of braided or gravel-bed rivers that have less clear channel boundaries than typical meandering rivers (Winterbottom, 2000).

Our analyses of georeferencing uncertainty support previous research recommending the use of second-order polynomials to optimize the combination of retained measurements and reduced image warping. We found significant differences in georeferencing uncertainty for images predating the 1990s (Fig. 8), likely due to reduced image quality and fewer reliable control points to georeference. An absence of local and global autocorrelation for GCPs reinforced our other results, identifying the need for spatially variable LoDs (SV-LoD) because errors associated with GCPs were uncorrelated irrespective of the distance between them. Spatial autocorrelation for both delineation bias and measured migration rates were autocorrelated from 1 to 6 channel widths (50- to 400-m). The scale of autocorrelation arising from coherent reaches of similar migration will likely vary with river size and are not directly transferrable to other systems. Thus, autocorrelation should be explored in future studies, and additional exploration may improve models of river meander migration.

Our analyses demonstrate how spatially-variable LoDs improve the quality of retained measurements relative to a singular RMSE-based LoD for two reasons. First, SV-LoDs are able to detect and eliminate erroneous large-magnitude migration values (false-positives) that arise from georeferencing offset or image warping in cases where the local spatially variable uncertainty is equally large (Fig. 7). In contrast, RMSE-based LoDs retain such measurements simply due to their large magnitude. Second, SV-LoDs retain the abundance of small, but legitimate, channel adjustments that often fall below the RMSE LoD (Fig. 8). Thus, we recommend the use of a SV-LoD in order to more accurately quantify uncertainty, as well as improve the quantity and quality of retained measurements. Currently, using the Planform Statistics Toolbox (Lauer and Parker, 2008) in combination with Lea and Legleiter's (2016) Matlab script quantifying spatially variable error provides an accurate, efficient, and nearly seamless means to quantify linear river migration and spatially variable uncertainty. For braided and/or anabranching channels, or rivers with many permanent vegetated islands, polygonal (area-based) methods of quantifying changes and associated spatially variable estimates (Gaeuman et al., 2005b) may be more suitable.

After applying a LoD threshold for parsing significant and nondetect

measurements, a few approaches may be appropriate for handling nondetects, contingent upon expert knowledge and scope/goals of the research (Fig. 12). In both linear and areal measurements of channel change, observations of 'zero' or nearly-zero change are generally flagged as 'nondetects', despite the possibility that no change actually occurred. In such cases, we recommend using expert discretion to discern if these measurements qualify as 'significant' measurements of zero, or nearly zero, change. Because the majority of river channels exhibit negligible adjustments between two photos, this exercise is likely to improve the accuracy of the final data distribution, as it reduces the odds of erroneously filtering measurements of real geomorphic change (Anderson, 2018). This recommendation is not justification for determining the significance of any/all measurements, but merely for evaluating nondetect measurements along known stagnant reaches. We evaluated the ability of three new approaches (i.e., Kaplan-Meier, Regression on Order Statistics, and Maximum Likelihood Estimation) to estimate statistical parameters (mean, median, standard deviation, and distribution fit) for modelled distributions with known proportions of nondetects (Table 3). MLE and K-M consistently perform well for approximating the mean of raw data at small measurement sample sizes ($n = 100$). However, at sample sizes > 1000, MLE will best approximate the mean. ROS will perform best for estimating the variance at all sample sizes and exhibits improvements in median estimates as sample size increases. KM is consistently the most robust in its overall distribution fit. The specific approach chosen for handling nondetects remains contingent upon each case, but should be guided and informed by the descriptions of each method and their requirements (Section 2.3.5), references to external resources, and results of our analyses.

Herein, we provided a comprehensive summary and evaluation of research on uncertainty as applied to studies of river planform change. Decades of research have built our understanding of uncertainty in remotely-sensed data, and will undoubtedly continue to be refined with improved technologies, software, statistical approaches, and most importantly, critical thinking. Future work should aim to improve upon the guidance provided herein to improve accuracy and uncertainty in measurements of fluvial change. There has been little consensus on applying knowledge gleaned from over two decades of research on identifying and quantifying uncertainty. This likely reflects the complicated nature of calculating uncertainty, the variety of tools (and thus, output formats) used to evaluate planform change, and in some cases, the absence of uncertainty estimates. Thus, we encourage improving the simplicity, generalizability, and open-source opportunities of tools and packages used for calculating planform change and associated uncertainty, thereby enabling a common platform to measure and compare results.

Acknowledgements

In addition to the three helpful reviews we received, we thank Sara Null, Devin Lea, and Lee Helsel for input and feedback in developing the methods and statistics along the way. Additional ideas and assistance with bank digitization were provided by Shannon Belmont, Devon Libby, Adam Fisher and Susan Durham.

Funding

This work was supported by the Quinney Foundation, National Science Foundation (NSF ENG 1209445), Minnesota Department of Agriculture, and Utah State University. The work also received support from the Utah Agricultural Experiment Station, approved as paper #9193.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.earscirev.2019.04.009.

References

Aalto, R., Lauer, J.W., Dietrich, W.E., 2008. Spatial and temporal dynamics of sediment accumulation and exchange along Strickland River floodplains (Papua New Guinea) over decadal-to-centennial timescales. J. Geophys. Res. 113. https://doi.org/10.1029/2006JF000627.

Allmendinger, N.E., Pizzuto, J.E., Moglen, G.E., Lewicki, M., 2007. A sediment budget for an urbanizing watershed, 1951-1996, Montgomery County, Maryland, U.S.A. J. Am. Water Resour. Assoc. 43, 1483–1498.

Anders, F.J., Byrnes, M., 1991. Accuracy of shoreline change rates as determined from maps and aerial photographs. Shore Beach 59, 17–26.

Anderson, S., 2018. Uncertainty in quantitative analyses of topographic change: error propagation and the role of thresholding: uncertainty in repeat topographic analyses. Earth Surf. Process. Landf. https://doi.org/10.1002/esp.4551.

Bailey, R.G., 1988. Problems with using overlay mapping for planning and their implications for geographic information systems. Environ. Manag. 12, 11–17. https://doi.org/10.1007/BF01867373.

Baki, A.B.M., Gan, T.Y., 2012. Riverbank migration and island dynamics of the braided Jamuna River of the Ganges–Brahmaputra basin using multi-temporal Landsat images. Quat. Int. 263, 148–161. https://doi.org/10.1016/j.quaint.2012.03.016.

Bangen, S., Hensleigh, J., McHugh, P., Wheaton, J., 2016. Error modeling of DEMs from topographic surveys of rivers using fuzzy inference systems. Water Resour. Res. 52, 1176–1193. https://doi.org/10.1002/2015WR018299.

Belmont, P., Gran, K.B., Schottler, S.P., Wilcock, P.R., Day, S.S., Jennings, C., Lauer, J.W., Viparelli, E., Willenbring, J.K., Engstrom, D.R., Parker, G., 2011. Large shift in source of fine sediment in the Upper Mississippi River. Environ. Sci. Technol. 45, 8804–8810. https://doi.org/10.1021/es2019109.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300.

Bjerklie, D.M., Moller, D., Smith, L.C., Dingman, S.L., 2005. Estimating discharge in rivers using remotely sensed hydraulic information. J. Hydrol. 309, 191–209. https://doi.org/10.1016/j.jhydrol.2004.11.022.

Blundell, B.J.S., Opitz, D.W., 2006. Object recognition and feature extraction from imagery: the Feature Analyst® approach (2006). In: ISPRS Proceedings. Presented at the International Conference on Object-Based Image Analysis (OBIA 2006). Salzburg Unversity, Austria.

Brasington, J., Langham, J., Rumsby, B., 2003. Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport. Geomorphology 53, 299–316. https://doi.org/10.1016/S0169-555X(02)00320-3.

Chrisman, N.R., 1987. The accuracy of map overlays: a reassessment. Landsc. Urban Plan. 14, 427–439. https://doi.org/10.1016/0169-2046(87)90054-5.

Clarke, T.A., Fryer, J.G., 1998. The development of camera calibration methods and models. Photogramm. Rec. 16, 51–66. https://doi.org/10.1111/0031-868X.00113.

Constantine, J.A., Dunne, T., Ahmed, J., Legleiter, C., Lazarus, E.D., 2014. Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. Nat. Geosci. 7, 899–903. https://doi.org/10.1038/ngeo2282.

Crosetto, M., Tarantola, S., 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. Int. J. Geogr. Inf. Sci. 15, 415–437. https://doi.org/10.1080/13658810110053125.

Dean, D.J., Schmidt, J.C., 2011. The role of feedback mechanisms in historic channel changes of the lower Rio Grande in the big Bend region. Geomorphology 126, 333–349. https://doi.org/10.1016/j.geomorph.2010.03.009.

Donovan, M., Belmont, P., 2019. Timescale dependence in river channel migration measurements. Earth Surf. Process. Landf. https://doi.org/10.1002/esp.4590.

Donovan, M., Miller, A., Baker, M., Gellis, A., 2015. Sediment contributions from floodplains and legacy sediments to Piedmont streams of Baltimore County, Maryland. Geomorphology 235, 88–105. https://doi.org/10.1016/j.geomorph.2015.01.025.

Donovan, M., Miller, A., Baker, M., 2016. Reassessing the role of milldams in Piedmont floodplain development and remobilization. Geomorphology 268, 133–145. https://doi.org/10.1016/j.geomorph.2016.06.007.

Downs, P.W., Dusterhoff, S.R., Sears, W.A., 2013. Reach-scale channel sensitivity to multiple human activities and natural events: lower Santa Clara River, California, USA. Geomorphology 189, 121–134. https://doi.org/10.1016/j.geomorph.2013.01.023.

Downward, S.R., Gurnell, A.M., Brookes, A., 1994. A methodology for quantifying river channel planform change using GIS. In: Proceedings of the Canberra Symposium. vol. 224. IAHS Publication, pp. 449–456.

Draut, A.E., Logan, J.B., McCoy, R.E., McHenry, M., Warrick, J.A., 2008. Channel evolution on the lower Elwha River, Washington, 1939-2006 (Report no. 2008–5127). In: Scientific Investigations Report. U.S. Geological Survey, Washington.

Edwards, G., Lowell, K.E., 1996. Modeling uncertainty in photointerpreted boundaries. Photogramm. Eng. Remote Sens. 62, 377–390.

Fay, M.P., Proschan, M.A., 2010. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. Stat. Surv. 4, 1–39. https://doi.org/10.1214/09-SS051.

Fisher, G.B., Bookhagen, B., Amos, C.B., 2013. Channel planform geometry and slopes from freely available high-spatial resolution imagery and DEM fusion: implications for channel width scalings, erosion proxies, and fluvial signatures in tectonically active landscapes. Geomorphology 194, 46–56. https://doi.org/10.1016/j.geomorph.2013.04.011.

Fraley, L.M., Miller, A.J., Welty, C., 2009. Contribution of in-channel processes to sediment yield of an urbanizing watershed. JAWRA J. Am. Water Resour. Assoc. 45, 748–766. https://doi.org/10.1111/j.1752-1688.2009.00320.x.

Fryer, J.G., Brown, D.C., 1986. Lens distortion for close-range photogrammetry. Photogramm. Eng. Remote Sens. 52, 51–58.
Gaeuman, D.A., Schmidt, J.C., Wilcock, P.R., 2003. Evaluation of in-channel gravel storage with morphology-based gravel budgets developed from planimetric data. J. Geophys. Res. 108https://doi.org/10.1029/2002JF000002. n/a-n/a.
Gaeuman, D., Schmidt, J.C., Wilcock, P.R., 2005a. Complex channel responses to changes in stream flow and sediment supply on the lower Duchesne River, Utah. Geomorphology 64, 185–206. https://doi.org/10.1016/j.geomorph.2004.06.007.
Gaeuman, D., Symanzik, J., Schmidt, J.C., 2005b. A map overlay error model based on boundary geometry. Geogr. Anal. 37, 350–369. https://doi.org/10.1111/j.1538-4632.2005.00585.x.
Geary, R.C., 1954. The contiguity ratio and statistical mapping. Inc. Stat. 5, 115. https://doi.org/10.2307/2986645.
Güneralp, İ., Rhoads, B.L., 2008. Continuous characterization of the planform geometry and curvature of meandering rivers: planform geometry and curvature of meandering rivers. Geogr. Anal. 40, 1–25. https://doi.org/10.1111/j.0016-7363.2007.00711.x.
Güneralp, İ., Filippi, A.M., Hales, B.U., 2013. River-flow boundary delineation from digital aerial photography and ancillary images using support vector machines. GIScience Remote Sens. 50, 1–25.
Güneralp, İ., Filippi, A.M., Hales, B., 2014. Influence of river channel morphology and bank characteristics on water surface boundary delineation using high-resolution passive remote sensing and template matching. Earth Surf. Process. Landf. 39, 977–986. https://doi.org/10.1002/esp.3560.
Gurnell, A.M., Downward, S.R., Jones, R., 1994. Channel planform change on the river Dee meanders, 1876–1992. Regul. Rivers Res. Manag. 9, 187–204. https://doi.org/10.1002/rrr.3450090402.
Harpold, A.A., Marshall, J.A., Lyon, S.W., Barnhart, T.B., Fisher, B., Donovan, M., Brubaker, K.M., Crosby, C.J., Glenn, N.F., Glennie, C.L., Kirchner, P.B., Lam, N., Mankoff, K.D., McCreight, J.L., Molotch, N.P., Musselman, K.N., Pelletier, J., Russo, T., Sangireddy, H., Sjöberg, Y., Swetnam, T., West, N., 2015. Laser vision: lidar as a transformative tool to advance critical zone science. Hydrol. Earth Syst. Sci. Discuss. 12, 1017–1058. https://doi.org/10.5194/hessd-12-1017-2015.
Helsel, D.R., 2005. More than obvious: better methods for interpreting nondetect data. Environ. Sci. Technol. 39, 419A–423A. https://doi.org/10.1021/es053368a.
Helsel, D.R., 2006. Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. Chemosphere 65, 2434–2439. https://doi.org/10.1016/j.chemosphere.2006.04.051.
Hickin, E., Nanson, G., 1984. Lateral migration rates of river bends. J. Hydraul. Eng. 110, 1557–1567. https://doi.org/10.1061/(ASCE)0733-9429(1984)110:11(1557).
Hooke, J.M., 1980. Magnitude and distribution of rates of river bank erosion. Earth Surf. Process. 5, 143–157. https://doi.org/10.1002/esp.3760050205.
Hosmer, D.W., Lemeshow, S., May, S., 2008. Applied survival analysis: regression modeling of time-to-event data. In: Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA. https://doi.org/10.1002/9780470258019.
Hossain, M.A., Gan, T.Y., Baki, A.B.M., 2013. Assessing morphological changes of the Ganges River using satellite images. Quat. Int. 304, 142–155. https://doi.org/10.1016/j.quaint.2013.03.028.
Hughes, M.L., McDowell, P.F., Marcus, W.A., 2006. Accuracy assessment of georectified aerial photographs: Implications for measuring lateral channel movement in a GIS. Geomorphology 74, 1–16. https://doi.org/10.1016/j.geomorph.2005.07.001.
Huston, C., Juarez-Colunga, E., 2009. Guidelines for computing summary statistics for data-sets containing non-detects. In: Statistical Guidance. Department of Statistics and Actuarial Science, Simon Fraser University, British Columbia, CA.
ITRC, 2013. Groundwater Statistics and monitoring Compliance. In: Statistical Tools for the Project Life Cycle. ITRC (Nondetects 5.7).
Kelly, S., Belmont, P., 2018. High resolution monitoring of river Bluff Erosion reveals failure mechanisms and geomorphically effective flows. Water 10, 394. https://doi.org/10.3390/w10040394.
Kiiveri, H.T., 1997. Assessing, representing and transmitting positional uncertainty in maps. Int. J. Geogr. Inf. Sci. 11, 33–52. https://doi.org/10.1080/136588197242482.
Ladd, G.B., Nagchaudhuri, A., Earl, T.J., Mitra, M., Bland, G.L., 2006. Rectification, georeferencing, and mosaicking of images acquired with remotely operated aerial platforms. In: Presented at the ASPRS 2006 Annual Conference, ASPRS, Reno, NV, pp. 10.
Lauer, J.W., 2007. Channel Planform Statistics Toolbox. National Center for Earth-Surface Dynamics, University of Minnesota.
Lauer, J.W., Parker, G., 2008. Net local removal of floodplain sediment by river meander migration. Geomorphology 96, 123–149. https://doi.org/10.1016/j.geomorph.2007.08.003.
Lauer, J.W., Echterling, C., Lenhart, C., Belmont, P., Rausch, R., 2017. Air-photo based change in channel width in the Minnesota River basin: modes of adjustment and implications for sediment budget. Geomorphology 297, 170–184. https://doi.org/10.1016/j.geomorph.2017.09.005.
Lawler, D.M., 1993. The measurement of river bank erosion and lateral channel change: a review. Earth Surf. Process. Landf. 18, 777–821.
Lea, D.M., Legleiter, C.J., 2016. Refining measurements of lateral channel movement from image time series by quantifying spatial variations in registration error. Geomorphology 258, 11–20. https://doi.org/10.1016/j.geomorph.2016.01.009.
Lee, L., Helsel, D., 2005. Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. Comput. Geosci. 31, 1241–1248. https://doi.org/10.1016/j.cageo.2005.03.012.
Legleiter, C.J., Kyriakidis, P.C., 2007. Forward and inverse transformations between cartesian and channel-fitted coordinate systems for Meandering Rivers. Math. Geol. 38, 927–958. https://doi.org/10.1007/s11004-006-9056-6.
Libby, D.J., Larson, P.H., Howell, D.W., Aeikens, A.L., Scheeler, D.J., Millett, J.J., Williams, V.A., Hilgendorf, Z., Chadwick-Camp, M., 2016. Assessing Error and Uncertainty in Remote Analysis of Channel Change Dynamics and Morphology. Case Study. (Minnesota River, Minnesota, USA).
Liro, M., 2015. Estimation of the impact of the aerialphoto scale and the measurement scale on the error in digitization of a river bank. Z. Geomorphol. 59, 443–453. https://doi.org/10.1127/zfg/2014/0164.
Martín-Fernández, J.A., Barceló-Vidal, C., Pawlowsky-Glahn, V., 2003. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Math. Geol. 35, 253–278. https://doi.org/10.1023/A:1023866030544.
Massey Jr., F.J., 1951. The kolmogorov-smirnov test for goodness of fit. J. Am. Stat. Assoc. 46, 68–78. https://doi.org/10.2307/2280095.
McMillan, M., Hu, Z., 2017. A watershed scale spatially-distributed model for streambank erosion rate driven by channel curvature. Geomorphology. https://doi.org/10.1016/j.geomorph.2017.03.017.
Melville, J.K., Martz, L.W., 2004. A comparison of data sources for manual and automated hydrographical network delineation. Can. Water Resour. J. 29, 267–282. https://doi.org/10.4296/cwrj267.
Micheli, E.R., Kirchner, J.W., 2002. Effects of wet meadow riparian vegetation on streambank erosion. 1. Remote sensing measurements of streambank migration and erodibility. Earth Surf. Process. Landf. 27, 627–639. https://doi.org/10.1002/esp.338.
Micheli, E.R., Kirchner, J.W., Larsen, E.W., 2004. Quantifying the effect of riparian forest versus agricultural vegetation on river meander migration rates, Central Sacramento River, California, USA. River Res. Appl. 20, 537–548. https://doi.org/10.1002/rra.756.
Miller, J.R., Friedman, J.M., 2009. Influence of flow variability on floodplain formation and destruction, Little Missouri River, North Dakota. Geol. Soc. Am. Bull. 121, 752–759. https://doi.org/10.1130/B26355.1.
Morais, E.S., Rocha, P.C., Hooke, J., 2016. Spatiotemporal variations in channel changes caused by cumulative factors in a meandering river: the lower Peixe River, Brazil. Geomorphology 273, 348–360. https://doi.org/10.1016/j.geomorph.2016.07.026.
Mount, N., Louis, J., 2005. Estimation and propagation of error in measurements of river channel movement from aerial imagery. Earth Surf. Process. Landf. 30, 635–643. https://doi.org/10.1002/esp.1172.
Mount, N., Louis, J., Teeuw, R., Zukowskyj, P., Stott, T., 2003. Estimation of error in bankfull width comparisons from temporally sequenced raw and corrected aerial photographs. Geomorphology 56, 65–77. https://doi.org/10.1016/S0169-555X(03)00046-1.
Mount, N.J., Tate, N.J., Sarker, M.H., Thorne, C.R., 2013. Evolutionary, multi-scale analysis of river bank line retreat using continuous wavelet transforms: Jamuna River, Bangladesh. Geomorphology 183, 82–95. https://doi.org/10.1016/j.geomorph.2012.07.017.
Nelson, N.C., Erwin, S.O., Schmidt, J.C., 2013. Spatial and temporal patterns in channel change on the Snake River downstream from Jackson Lake dam, Wyoming. Geomorphology 200, 132–142. https://doi.org/10.1016/j.geomorph.2013.03.019.
Osterkamp, W.R., Hedman, E.R., 1982. Perennial-streamflow characteristics related to channel geometry and sediment in Missouri River basin (Report no. 1242). In: Professional Paper.
Passalacqua, P., Belmont, P., Staley, D.M., Simley, J.D., Arrowsmith, J.R., Bode, C.A., Crosby, C., DeLong, S.B., Glenn, N.F., Kelly, S.A., Lague, D., Sangireddy, H., Schaffrath, K., Tarboton, D.G., Wasklewicz, T., Wheaton, J.M., 2015. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: a review. Earth Sci. Rev. 148, 174–193. https://doi.org/10.1016/j.earscirev.2015.05.012.
Pavelsky, T.M., Smith, L.C., 2008. RivWidth: a software tool for the calculation of river widths from remotely sensed imagery. IEEE Geosci. Remote Sens. Lett. 5, 70–73. https://doi.org/10.1109/LGRS.2007.908305.
Peixoto, J.M.A., Nelson, B.W., Wittmann, F., 2009. Spatial and temporal dynamics of river channel migration and vegetation in central Amazonian white-water floodplains by remote-sensing techniques. Remote Sens. Environ. 113, 2258–2266. https://doi.org/10.1016/j.rse.2009.06.015.
Piégay, H., Cuaz, M., Javelle, E., Mandier, P., 1997. Bank erosion management based on geomorphological, ecological and economic criteria on the Galaure River, France. Regul. Rivers Res. Manag. 13, 433–448. https://doi.org/10.1002/(SICI)1099-1646(199709/10)13:5<433::AID-RRR467>3.0.CO;2-L.
Piégay, H., Darby, S.E., Mosselman, E., Surian, N., 2005. A review of techniques available for delimiting the erodible river corridor: a sustainable approach to managing bank erosion. River Res. Appl. 21, 773–789. https://doi.org/10.1002/rra.881.
Rhoades, E.L., O'Neal, M.A., Pizzuto, J.E., 2009. Quantifying bank erosion on the South River from 1937 to 2005, and its importance in assessing Hg contamination. Appl. Geogr. 29, 125–134. https://doi.org/10.1016/j.apgeog.2008.08.005.
Rowland, J.C., Shelef, E., Pope, P.A., Muss, J., Gangodagamage, C., Brumby, S.P., Wilson, C.J., 2016. A morphology independent methodology for quantifying planview river change and characteristics from remotely sensed imagery. Remote Sens. Environ. 184, 212–228. https://doi.org/10.1016/j.rse.2016.07.005.
Schaffrath, K.R., Belmont, P., Wheaton, J.M., 2015. Landscape-scale geomorphic change detection: quantifying spatially variable uncertainty and circumventing legacy data issues. Geomorphology 250, 334–348. https://doi.org/10.1016/j.geomorph.2015.09.020.
Schook, D.M., Rathburn, S.L., Friedman, J.M., Wolf, J.M., 2017. A 184-year record of river meander migration from tree rings, aerial imagery, and cross sections. Geomorphology 293, 227–239. https://doi.org/10.1016/j.geomorph.2017.06.001.
Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V., Foufoula-Georgiou, E., 2017. High spatiotemporal resolution of river planform dynamics from Landsat: the RivMAP toolbox and results from the Ucayali River: annual planform morphodynamics, ucayali. Earth Space Sci. 4, 46–75. https://doi.org/10.1002/2016EA000196.
Shields, F.D., Simon, A., Steffen, L.J., 2000. Reservoir effects on downstream river

channel migration. Environ. Conserv. 27, 54–66.
Shumway, R.H., Azari, R.S., Kayhanian, M., 2002. Statistical approaches to estimating mean water quality concentrations with detection limits. Environ. Sci. Technol. 36, 3345–3353. https://doi.org/10.1021/es0111129.
Singh, A., Nocerino, J., 2002. Robust estimation of mean and variance using environmental data sets with below detection limit observations. Chemom. Intell. Lab. Syst. 60, 69–86. https://doi.org/10.1016/S0169-7439(01)00186-1.
Smith, L.C., Pavelsky, T.M., 2008. Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia: river discharge and hydraulic geometry. Water Resour. Res. 44. https://doi.org/10.1029/2007WR006133.
Smith, S., Belmont, P., Wilcock, P.R., 2011. Closing the gap between watershed modeling, sediment budgeting, and stream restoration. In: Simon, A., Bennett, S.J., Castro, J.M. (Eds.), Geophysical Monograph Series. Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools American Geophysical Union, Washington, D. C, pp. 293–317.
Souffront, M., 2014. Channel adjustment and channel-floodplain sediment exchange in the root River, Southeastern Minnesota (M.S. Thesis). Utah state University, Logan, Utah.
Spiekermann, R., Betts, H., Dymond, J., Basher, L., 2017. Volumetric measurement of river bank erosion from sequential historical aerial photography. Geomorphology 296, 193–208. https://doi.org/10.1016/j.geomorph.2017.08.047.
Swanson, B.J., Meyer, G.A., Coonrod, J.E., 2011. Historical channel narrowing along the Rio Grande near Albuquerque, New Mexico in response to peak discharge reductions and engineering: magnitude and uncertainty of change from air photo measurements. Earth Surf. Process. Landf. 36, 885–900. https://doi.org/10.1002/esp.2119.
Sylvester, Z., Durkin, P., Covault, J.A., 2019. High curvatures drive river meandering. Geology 47, 263–266. https://doi.org/10.1130/G45608.1.
Tauber, F., 1999. Spurious Clusters in Granulometric Data Caused by Logratio Transformation. Math. Geol. 31, 491–504. https://doi.org/10.1023/A:1007532222145.
Unwin, D.J., 1995. Geographical information systems and the problem of "error and uncertainty". Prog. Hum. Geogr. 19, 549–558. https://doi.org/10.1177/030913259501900408.
USDA FSA APFO, 2015. National Agricultural Imagery Program (NAIP). Digital Orthorectified Images (DOQ), Minnesota.
Vericat, D., Wheaton, J.M., Brasington, J., 2017. Revisiting the Morphological Approach: Opportunities and challenges with repeat High-Resolution Topography. In: Tsutsumi, D., Laronne, J. (Eds.), Gravel-Bed Rivers: Processes and Disasters, Gravel Bed Rivers. John Wiley and Sons Ltd., pp. 121–155.
Ward, J.V., Tockner, K., Arscott, D.B., Claret, C., 2002. Riverine landscape diversity. Freshw. Biol. 47, 517–539. https://doi.org/10.1046/j.1365-2427.2002.00893.x.
Werbylo, K.L., Farnsworth, J.M., Baasch, D.M., Farrell, P.D., 2017. Investigating the accuracy of photointerpreted unvegetated channel widths in a braided river system: a Platte River case study. Geomorphology 278, 163–170. https://doi.org/10.1016/j.geomorph.2016.11.003.
Wheaton, J.M., Brasington, J., Darby, S.E., Sear, D.A., 2010. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. Earth Surf. Process. Landf. 35, 136–156. https://doi.org/10.1002/esp.1886.
Winterbottom, S.J., 2000. Medium and short-term channel planform changes on the Rivers Tay and Tummel, Scotland. Geomorphology 34, 195–208. https://doi.org/10.1016/S0169-555X(00)00007-6.